A SPATIAL ECONOMETRIC ESTIMATION MODEL

FOR U.S. FARMLAND VALUES

A Thesis

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Master of Science

by

Yingzong Sun

January 2014

ABSTRACT

This study investigates the economic problem of recent booming farmland values. An income capitalization model is estimated to conduct farmland valuations using state-level data from 1980 to 2011. Explanatory variables include expected market returns, government payments, production risk, urban influence, interest rates, and ethanol production scale. Spatial models are introduced to control for spatial dependencies on farmland values, and multiple tests are conducted to explore the most appropriate model for farmland valuation. Furthermore, the thesis offered suggestions for future researches and to provide a proposal in forecasting future farmland values according to the changes in the determining factors.

BIOGRAPHICAL SKETCH

Yingzong Sun was born March 15th, 1988 in Shenyang, China. He attended South China University of Technology in September 2006, and in 2009 he transferred to Purdue University, pursuing in the major of agribusiness management. Yingzong received his Bachelor of Science from Purdue University in 2011.

During his graduate studies in Cornell University, Yingzong worked as a research assistant and focused on the study of U.S. farmland valuations.

给我最亲爱的家人们

ACKNOWLEDGEMENTS

First of all, I want to express my deepest gratitude to my committee chair, Professor Joshua Woodard. His support consistently exceeded my expectations during the past two years. Without his advice, encouragement, patience and persistent help, the completion of this thesis would not have been possible.

I want to express my sincere thanks to my committee member, Professor Calum Turvey, for his valuable advice and continued guidance extended to me. I also want to thank Cornell Institute for Social and Economic Research for their assistance in technical aspects of the research.

And finally, I want to thank all my family members for their constant encouragement and greatest supports, during the course of past two years.

TABLE OF CONTENTS

1. Introduction

Being the most important asset on the farm balance sheet, farmland accounts for approximately 85% of total asset values in the U.S. agricultural sector (Henderson et al., 2011). Market participants are interested in farmland values not only because of the significance of land in agricultural production, but also because the changes in land values serve as an important signal of farm financial stress. During recent years, farmland values have increased dramatically, and many investors have turned to farmland as an alternative investment outlet. Therefore, an increasing number of economists have become interested in how the changes in farmland values might affect the health of the U.S. agricultural industry in the future.

1.1 Economic Problem

Subsequent to the farmland crisis of the 1980s, farmland values grew at a relatively stable rate for some time, but after 2005 growth in land values began to accelerate. According to data from the Land Values 2012 Summary (USDA, 2012), the average U.S. farmland price has doubled during the past decade, reaching record-highs in recent years (Figure 1.1).

The increase in farmland values is due to various agricultural and economic factors. Henderson et al. (2011), who studied agricultural boom and bust cycles in the United States, argue that the skyrocketing farmland prices today are mainly fueled by booming agricultural incomes. In addition, Henderson (2008) implied lean supplies, strong agricultural export activities, and increased demand from ethanol have all contributed to the boost of agricultural productions in the U.S. as well.

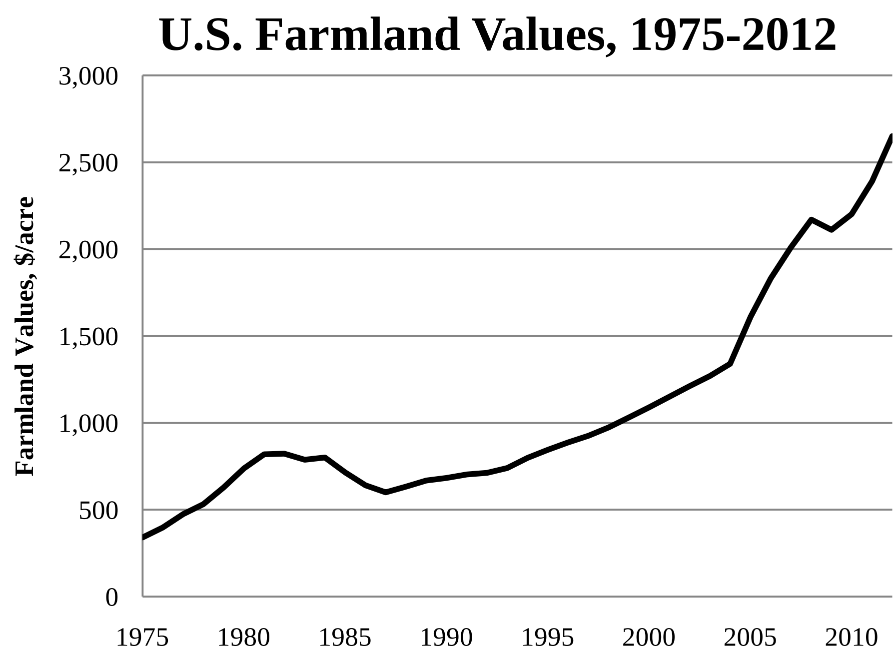

Figure 1.1 U.S. Farmland Values from 1975 to 2011, Measured on January 1st of Each Year. (Data Source: USDA-National Agricultural Statistics Service, Agricultural Statistics, Agricultural Land-Asset Value)

Despite large relative fluctuations from year to year, long-run price levels for the major field crops in the U.S. had been relatively stable since the mid-1970's, but have increased largest rapidly since about 2005 (Figure 1.2). For example, corn prices remained in the $2-$3 per bushel range for most of the last few decades, but increased precipitously after 2005, reaching into the $6/bushel range in recent years.

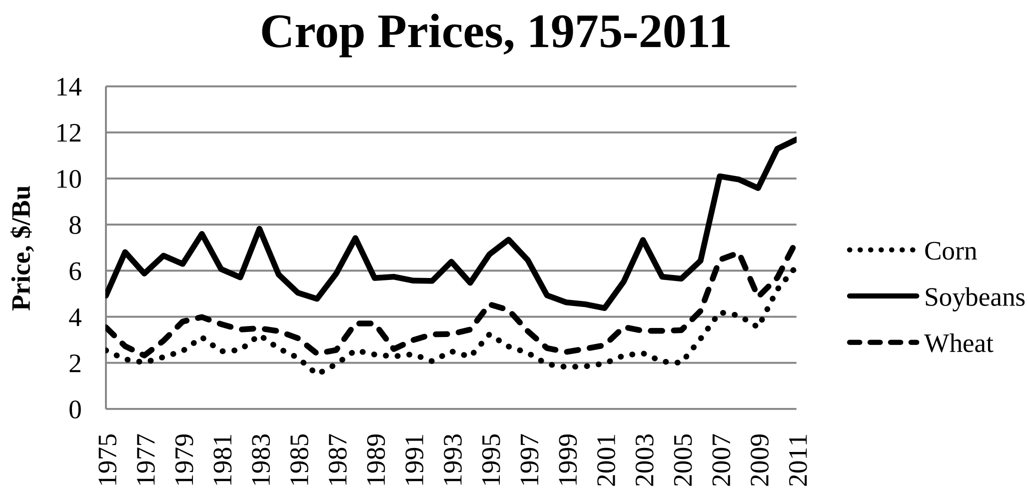

Figure 1.2 Crop Cash Prices, Measured as Year Average from 1975-2011. (Data Source: USDA-National Agricultural Statistics Service, Agricultural Statistics, Field Crops-Price Received)

The escalating farmland price begs the question of how farmland prices might change in the future. Gloy (2012) argues that the dramatic increases in farmland prices are less likely to be sustainable. By introducing a number of economic fundamentals as farmland valuation metrics, Gloy provided a clue to examine if the farmland market is rationally corresponding to economic conditions. According to his study, the farmland value-to-rent multiple reached record highs during the last decade, meanwhile, the rate of returns to productive farmland remained at historically low levels. These fundamentals suggest that the farmland prices are extremely high relative to income generation. Gloy points out that the interest rate declines during the past two decades, which combined with increasing farm incomes has contributed to escalating farmland values.

In addition, Kropp and Peckham (2012) indicate the large scale expansion of ethanol production has also contributed to the rising farmland values in recent years. The increase in ethanol production has increased demand for corn, leading to higher corn prices, and thus incentivizing farmers to plant more acres of corn. As corn production increases, other crops will compete with corn for land, lowering supply and increasing prices which drive up demand for farmland and increases farmland price.

Gloy (2012) and Henderson et al. (2011), who studied the recent boom of farmland values, argue that to keep the trend of increasing farmland values, investors have to be optimistic and expect farm incomes to keep growing while interest rates remain at current low levels. By observing historical economic fundamentals and comparing those with the current period, Gloy (2012) concluded that the increases in farmland values today are on par with the increases in the late 1970s.

1.2 Analytical Framework

Even though Gloy (2012) and Henderson et al. (2011) did not provide any quantitative modeling link between farmland values and the underlying fundamentals, a tremendous number of previous studies have developed various econometric methods for addressing this problem (for example, see Palmquist and Danielson, 1989; Xu et al., 1993; and Weersink et al., 1999). The capitalization model is a common pricing framework that economists often employ (Vyn et al., 2012; Kropp and Peckham, 2012). This model is based on the assumption that the value of a parcel of farmland should equal the sum of all discounted future cash flows associated with the land. Following Weersink et al. (1999) and Goodwin et al. (2003), in the capitalization model, market returns from agricultural production and payments from government programs are two important sources of income provided to land owners, and thus are always treated as key influential factors to farmland values. Production risk should also be considered because it measures the variability in future incomes (Vyn et al., 2012). As the majority of people are considered to be risk-adverse, higher production risk is expected to result in lower farmland values.

Moreover, a number of economic factors are always included in the capitalization model. Goodwin et al. (2003) and Huang (2006) indicate that urban pressure to large extent influences farmland values, since land prices should reflect not only the current use of land, but also the opportunity of converting land to alternative uses. High levels of urbanization increase the potential to convert farmland into residential or recreational uses, and this is likely to drive land values. In addition, interest rates are always introduced as an important economic factor because they affect cost of capital and investment activities.

As discussed above, ethanol production and the residual demand for corn exert great pressure on farmland values, and therefore ethanol production is introduced as another underlying determinant. Moreover, Kropp and Peckman (2012) point out that the effect of ethanol production varies across geographic locations; *ceteris paribus*, farmland located near ethanol facilities are expected to have higher price, because the ethanol facilities have greater impact on crop demands (especially corn demands) in local markets. To further explore the influence of ethanol production as well as to incorporate different levels of impact across space, our study uses multiple approaches and introduces a number of ethanol-space-interaction independent variables.

Furthermore, spatial dependencies of farmland values are addressed in our econometric model. In theory the basic ideas are presented in Hardie, Narayan, and Gardner (2001) who indicate that property values can be correlated across regions, in such a way that the price of a parcel of farmland is likely to be influenced by neighboring farmland values. Huang et al. (2006) argues that economic studies ignoring the effect of spatial dependencies can affect the consistency and efficiency of the estimates. By adopting a spatial lag model, they employ data in Illinois and explore the spatial effect of farmland prices. Their results showed the presence of spatial dependencies across Illinois counties, and demonstrated that the farmland price in a county will increase by 0.284% given a 1% land price increase in nearby counties.

Additionally, Woodard et al. (2010) use Illinois county-level and farm-level data from 1996 to 2008 to study the factors affecting agricultural cash rents. They incorporate spatial correlation into a basic hedonic model where the independent variables are separated into three categories: parcel characteristics, regional characteristics, and economic characteristics. The results demonstrated significant impacts on farmland rents from all categories, and they indicate that in order to better interpret data and study the influential determinants on farmland rents, spatial dependencies should be considered.

Following Huang et al. and Woodard et al., our study explicitly incorporates spatial dependence and correlation effects into the estimation of the econometric model, however our approach differs in a number of aspects. First, while both Huang et al. and Woodard et al. base their studies on hedonic pricing models, we adopt a capitalization model which mainly focuses on income factors and economic factors, rather than focusing on physical factors such as soil quality indexes. In addition, Huang et al. and Woodard et al. employ county-level and/or farm-level farmland values/rents data to examine the spatial dependencies within a small region (Illinois), whereas we are interested in spatial correlations across the entire U.S. To our knowledge, this is the first spatially explicit large scale model investigation of U.S. agricultural land values. Figures 1.3 and 1.4 depict state-level farmland values across the U.S. in 1980 and 2011, respectively. Strong spatial effects are apparent in both figures. This is not surprising given that the drivers of land values tend to be spatially correlated as well.

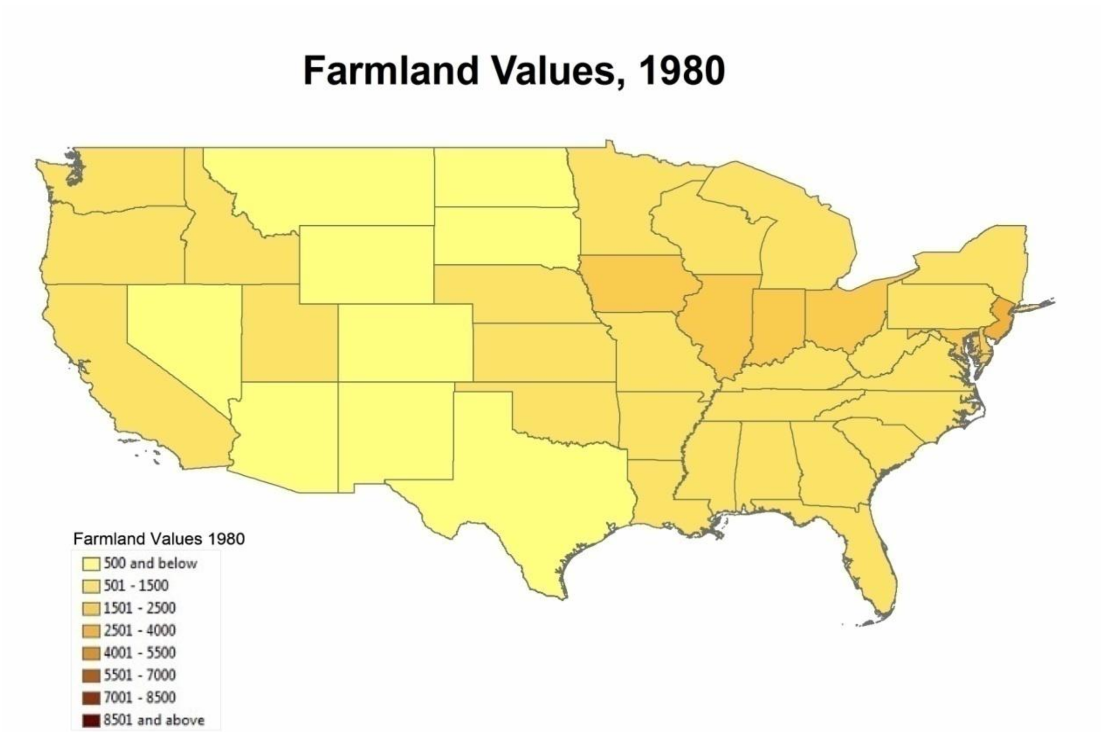

Figure 1.3 U.S. Farmland Values, Measured on January 1st, 1980 (Data Source: USDA-National Agricultural Statistics Service, Agricultural Statistics, Agricultural Land-Asset Value)

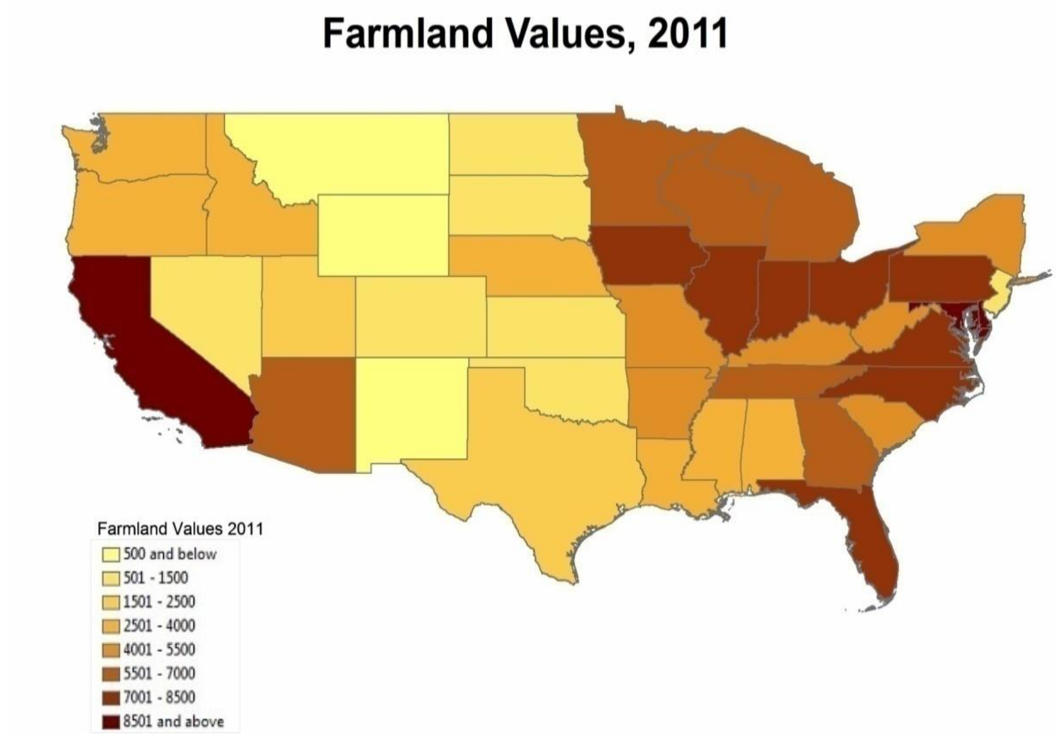

Figure 1.4 U.S. Farmland Values, Measured on January 1st, 2011 (Data Source: USDA-National Agricultural Statistics Service, Agricultural Statistics, Agricultural Land-Asset Value)

1.3 Purpose and Objectives

The purpose of our research is to investigate the economic problem of recent booming farmland values and using economic methods to conduct farmland valuations in the presence of rising agricultural incomes, low interest rates, large ethanol production increases, as well as other agricultural and economic factors. The specific objectives include:

1) To explore the most appropriate econometric model on farmland valuation.
2) To determine factors influencing farmland values and to examine the relationships between underlying determinants and farmland values.
3) To test spatial dependencies in farmland values across different states in the U.S.

4) To offer suggestions for future researches and to provide a proposal in forecasting future farmland values according to the changes in the determining factors.

2. Background

2.1 Conceptual Framework

As discussed above, a capitalization model is used in our study. The model assumes that the value of an asset is determined by a stream of discounted future cash returns. The framework can be specified in the equation as follows:

$$V_t = \sum_{j=1}^{\infty} \frac{E_t(R_{t+j})}{(1+r_{t+1})(1+r_{t+2})\ldots(1+r_{t+j})},$$

where V_t represents asset value at time *t*; R_t represents returns generated from the asset; E_t represents expectation based on information available in period *t*; and r_t represents discount rate. Hamilton and Whiteman (1985) indicate that, if it is assumed that discount rate remains constant over time and all agents are risk neutral, the model can be modified as:

$$V_t = \sum_{j=1}^{\infty} \frac{E_t(R_{t+j})}{(1+r)^j}$$

Furthermore, by assuming constant expected returns and infinite periods, Weersink et al. (1999) derived the model into the equation as follows:

$$V_t = \frac{E_t(R)}{r}$$

If it is assumed that the expected returns grow at a constant rate, *g*, the formula is transformed into:

$$V_t = \frac{E_t(R_{t+1})}{r-g}$$

These two models above have served as the basis of most of the asset valuations, including farmland. Following Weersink et al. (1999) and Vyn et al. (2012), the present value model of farmland valuation can be shown in the equation below:

$$L_t = \sum_{j=1}^{\infty} b^j E_t R_{t+j} \text{ ,}$$

where L_t is the farmland value at time *t*; R_t is the return generated during the period from *t* to *t+1*; E_t implies the expectation is made conditional to the information known at time *t*; and b^j is the discount factor. Due to the fact that returns can be generated from both the market (sales of crops) and government programs, several studies (Goodwin and Ortalo-Magne, 1992; Weersink et al., 1999; and Goodwin et al., 2003) disaggregate returns further into two components according to the income sources. Letting $M_{i,t+s}$ be the market return and $G_{i,t+s}$ be the government payments, the equation can be shown as:

$$L_t = \sum_{j=1}^{\infty} (b_1^j E_t M_{t+j} + b_2^j E_t G_{t+j}),$$

where b_1 and b_2 represent discount factors for market returns and government payments, respectively. Following Goodwin et al. (2003), the discount rates associated with market returns and government payments may differ from each other because different sources of returns have different level of risk. The discount rates from one source of returns may also change over time. However, if it is assumed that discount rates remain constant for the same source of returns (i.e., $b_1^j = b_1$ and $b_2^j = b_2$), and the cash flows grow at a constant rate, the formula can be expressed as,

$$L_t = \beta_1 E_t M_{t+1} + \beta_2 E_t G_{t+1}$$

Moreover, researchers such as Goodwin et al. (2003), Shaik et al. (2005), and Kropp and Peckham (2012) all incorporate various economic factors such as urban influence, locations, and population densities. The model is adjusted as:

$$L_t = \beta_1 E_t M_{t+1} + \beta_2 E_t G_{t+1} + \sum_{i=1}^{n} \beta_{3i} K_{i,t}$$

Where $K_{i,t}$ represents the *i*th economic factor, and β_{3i} represents the discount factor associated with it.

2.2 Instrumental Variables Approach

Following the concepts introduced in the last section, by introducing expected market returns (MR), expected government payments (GP), production risk ($Risk$), real interest rates ($Interest$), urban influence ($Urban$), and ethanol production ($Ethanol$) as explanatory variables, the traditional income capitalization model can be represented as:

$$L = f(MR, GP, Risk, Interest, Urban, Ethanol)$$

Even though this model has been widely used in the study of farmland values, a number of researchers such as Kropp and Peckham (2012), and Shaik et al. (2005) indicate that endogeneity issues may exist with the government payments variable, which may result in biased estimation results. This endogeneity may stem from people's expectations: for example, if people have optimistic expectations about future government payments, they are willing to pay higher price for farmland, and thus the farmland becomes more valuable. Following Shaik et al. (2005), to address the endogeneity issue, an instrumental variable approach is used in our study. An additional equation is introduced where the government payments serve as the dependent variable; predicted values for government payments from this equation are subsequently used to estimate the farmland value equation. The joint model can be expressed as:

$$\begin{aligned} L &= f(MR, GP, Risk, Interest, Urban, Ethanol) \\ GP &= f(MR, Urban, \text{Ethanol}, FB_j) \end{aligned},$$

where FB_j represents the *j*th farm bill period and serves as an instrument in our model.

2.3 Spatial Models

Spatial lag models and spatial error models are two common frameworks that economists adopt to incorporate spatial dependence (Anselin, 1988) Even though these approaches share great similarities in mathematical expressions, they differ in economic interpretation. The spatial lag model incorporates spatial effects pertaining to the dependent variable, whereas the spatial error model focuses on the spatial correlation in the error term only (Anselin, 1988). We will outline the structures for those two approaches and conduct our empirical analysis on farmland values using both.

2.3.1 Spatial Weight Matrix

Spatial weight matrices are employed in spatial econometrics to account for spatial effects. A spatial weight matrix, denoted by W_N, is an $N \times N$ matrix where N is the number of cross-sectional observations (in our study, the number of states). Elements in W_N

represent the strength of interaction between the corresponding states; for example, an element $w_{i,j}$ measures state *i* and *j*'s spatial impact on each other. For simplicity reasons, we suppose all neighbors of an observed state exert equal impact on that state, and non-neighbors have no influence on each other. In order to standardize the spatial effect, elements in each row of the weight matrix should sum to 1. Suppose state *i* has *m* neighbors, then $w_{i,j} = \frac{1}{m}$ when *i* and *j* are contiguous to each other, and $w_{i,j} = 0$ when they are not (this is referred to as a queen matrix, for details see Anselin, 1988). In our study, the weight matrix *W* is a sparse matrix; the majority of elements equal 0 because a given state only has a limited number of neighbors.

2.3.2 Spatial Lag Model

The spatial lag model assumes that correlation exists in the dependent variable (Huang, 2006). In the panel case, suppose *T* is the total number of time periods and *N* is the total number of states, the model can be expressed as follows:

$$L = \rho(I_T \otimes W_N)L + X\beta + \varepsilon,$$

where *L* is an $NT\times 1$ vector of farmland values per acre measured at state level; *X* is an $NT\times K$ matrix of independent variables; β is a $K\times 1$ vector of regression coefficients; ρ is the spatial autoregressive coefficient; I_T is an identity matrix with dimension *T*; W_N is an $N\times N$ spatial weight matrix; $\otimes$ is a sign for the Kronecker product (note that $I_T \otimes W_N$ yields a $NT \times NT$ block matrix, $I_T \otimes W_N = \begin{pmatrix} W_N & & 0 \\ & \ddots & \\ 0 & & W_N \end{pmatrix}_{NT\times NT}$); and ε is an

$NT \times 1$ vector of error terms.

Following Anselin (1988), the error term ε follows a normal distribution, ε ~ N(0, $\sigma^2 I$); and the spatial dependencies are captured in the matrix ($I_T \otimes W_N$). By converting the equation, we can express the function as:

$$L = [I_T \otimes (I_N - \rho W_N)^{-1}]X\beta + [I_T \otimes (I_N - \rho W_N)^{-1}]\varepsilon$$

Anselin (1988) shows OLS estimation may affect the consistency of the results and can cause biased estimates. Maximum likelihood estimators are commonly used in estimation of the spatial lag models. By maximizing the following likelihood function, one can estimate β, ρ, and σ^2 :

$$L = -\frac{[L - \rho(I_T \otimes W_N)L - X\beta]'[L - \rho(I_T \otimes W_N)L - X\beta]}{2\sigma^2} - \frac{N}{2}\ln(\frac{\pi}{2}) - \frac{N}{2}\ln(\sigma^2) + \ln[I_{NT} - \rho(I_T \otimes W_N)]$$

2.3.3 Spatial Error Model

The spatial error model assumes that there are omitted variables in the error term that are spatially correlated (Huang, 2006). A general form of the spatial error model can be expressed as:

$$L = X\beta + \varepsilon$$

Similar to the construction of spatial lag specification, the error vector ε can be modeled as:

$$\varepsilon = \rho(I_T \otimes W_N)\varepsilon + u,$$

where ρ is the spatial autoregressive parameter; I_T is an identity matrix; W_N is an spatial weight matrix; and u is a $N\times 1$ vector of reminder errors. The equation above can be expressed as:

$$\varepsilon = [I_T \otimes (I_N - \rho W_N)^{-1}]u \text{ ,}$$

The error component u is not spatially correlated because the spatial dependencies are captured via the spatial filter matrix. Anselin (1988) indicates that even though the OLS estimators are unbiased in the estimation of spatial error model, they can be inefficient; therefore the spatial estimators are preferred. By maximizing the following likelihood function, one can estimate β, ρ, and σ^2:

$$L = -\frac{[Y - X\beta]'[Y - \rho(I_T \otimes W_N)Y]'[Y - X\beta]}{\sigma^2} - \frac{N}{2}\ln(\frac{\pi}{2}) - \frac{N}{2}\ln(\sigma^2)$$
$$+\ln[I_{NT} - \rho(I_T \otimes W_N)]$$

3. Empirical Models

Our study starts with a basic capitalization model, which can be expressed as:

$$\text{(1)} \quad L_{i,t} = \beta_0 + \beta_1 MR_{i,t} + \beta_2 GP_{i,t} + \beta_3 Risk_{i,t} + \beta_4 Interest_t + \beta_5 Urban_{i,t} + \beta_6 Ethanol_t + \varepsilon_{i,t}$$

where the denotation is the same as used in last section; *i* and *t* represent cross sectional and time series dimensions, respectively. As discussed above, in order to account for the geographic impact of ethanol production on farmland values (note that the *Ethanol* variable represents annual ethanol production in the entire U.S.), we introduce two additional models as described below:

1) The first model uses the spatial expansion method (Cassetti, 1972), which

includes spatially varying factors represented by longitude and latitude coordinates of different states. The model can be expressed as:

$$
\begin{aligned}
L_{i,t} &= \beta_0 + \beta_1 MR_{i,t} + \beta_2 GP_{i,t} + \beta_3 Risk_{i,t} + \beta_4 Interest_t + \beta_5 Urban_{i,t} + \\
&(\beta_6 + \beta_7 Coordx + \beta_8 Coordy + \beta_9 Coordx * Coordy) * Ethanol_t + \varepsilon_{i,t}
\end{aligned} \tag{2}
$$

where $Coordx$ and $Coordy$ denote longitude and latitude coordinates, respectively. $Coordx * Ethanol$, $Coordy * Ethanol$, and $Coordx * Coordy * Ethanol$ represent the spatial interactions between locations and ethanol production.

2) The second model introduces a new spatial-interaction factor, $Dist_i$, which represents the distance between state i and the center of the Midwest area. Because the majority of ethanol plants in the U.S. are located in the Midwest area, we expect the ethanol productions would have greater impacts on farmland values in states closer to the Midwest. The model can be expressed as:

$$
\begin{aligned}
L_{i,t} &= \beta_0 + \beta_1 MR_{i,t} + \beta_2 GP_{i,t} + \beta_3 Risk_{i,t} + \beta_4 Interest_t + \beta_5 Urban_{i,t} \\
&+ (\beta_6 + \beta_7 Dist) * Ethanol_t + \varepsilon_{i,t}
\end{aligned} \tag{3}
$$

Specifications (1) through (3) serve as the basis of our capitalization model. Furthermore, we use an instrumental variables approach to address the endogeneity issues with the government payments variable. As discussed in the previous section, an additional equation is introduced where government payments are expressed as a function of market returns, urban influence, ethanol production, and time period dummy variables for seven farm bill periods. The joint model can be represented as:

$$(4)\quad \begin{aligned} L_{i,t} &= \beta_0 + \beta_1 MR_{i,t} + \beta_2 GP_{i,t} + \beta_3 Risk_{i,t} + \beta_4 Interest_t + \beta_5 Urban_{i,t} \\ &+(\beta_6 + \beta_7 Coordx + \beta_8 Coordy + \beta_9 Coordx * Coordy) * Ethanol_t + \varepsilon_{i,t} \end{aligned}$$

$$GP_{i,t} = \gamma_0 + \gamma_1 MR_{i,t} + \gamma_2 Urban_{i,t} + \gamma_3 Ethanol_t + \sum_{j=2}^{7} \gamma_{4,j} FB_j + +\delta_{i,t}$$

and

$$(5)\quad \begin{aligned} L_{i,t} &= \beta_0 + \beta_1 MR_{i,t} + \beta_2 GP_{i,t} + \beta_3 Risk_{i,t} + \beta_4 Interest_t + \beta_5 Urban_{i,t} \\ &+(\beta_6 + \beta_7 dist) * Ethanol_t + \varepsilon_{i,t} \end{aligned}$$

$$GP_{i,t} = \gamma_0 + \gamma_1 MR_{i,t} + \gamma_2 Urban_{i,t} + \gamma_3 Ethanol_t + \sum_{j=2}^{7} \gamma_{4,j} FB_j + +\delta_{i,t}$$

Moreover, following Huang et al. (2006) and Woodard et al. (2010), our research adopts a spatial model which incorporates spatial dependencies. Spatial lag and spatial error frameworks are investigated separately to account for the effect of spatial dependencies. Using the instrumental variables approach in these spatial frameworks, a government payment equation is estimated first, and the predicted values are used in place of expected government payments to estimate the spatial model. The models are expressed as follows:

Spatial Lag Model:

$$(6)\quad \begin{aligned} L_{i,t} &= \beta_0 + \rho(I_T \otimes W_N) L_{i,t} + \beta_1 MR_{i,t} + \beta_2 GP_{i,t} + \beta_3 Risk_{i,t} + \beta_4 Interest_t + \beta_5 Urban_{i,t} + \\ &\beta_6 Ethanol_t + \varepsilon_{i,t} \end{aligned}$$

$$GP_{i,t} = \gamma_0 + \gamma_1 MR_{i,t} + \gamma_2 Urban_{i,t} + \gamma_3 Ethanol_t + \sum_{j=2}^{7} \gamma_{4,j} FB_j + +\delta_{i,t}$$

Spatial Error Model:

$$(7)\quad L_{i,t} = \beta_0 + \beta_1 MR_{i,t} + \beta_2 GP_{i,t} + \beta_3 Risk_{i,t} + \beta_4 Interest_t + \beta_5 Urban_{i,t} + \beta_6 Ethanol_t + \varepsilon_{i,t}$$

$$\varepsilon_{i,t} = \rho(I_T \otimes W_N)\varepsilon_{i,t} + u_{i,t}$$

$$GP_{i,t} = \gamma_0 + \gamma_1 MR_{i,t} + \gamma_2 Urban_{i,t} + \gamma_3 Ethanol_t + \sum_{j=2}^{7} \gamma_{4,j} FB_j + + \delta_{i,t}$$

Various indicators have been used to account for expected market returns in previous studies, such as net farm income (Goodwin et al., 2003), crop returns (Shaik et al., 2005), and crop receipts and expenses (Vyn et al., 2012). In order to explore the most appropriate model to measure farmland values, we adopt two approaches and use different market return indicators: 1) net farm income from operations; 2) a crop profits index built by including various production factors, including detrended yields, expected crop prices, production scale, and costs.

4. Data

We obtain agricultural net cash incomes, government payments, and Beale Index (as an indicator of urban influence) data from USDA-Economic Research Service; farmland values and other farm production data including, crop production yields, crop prices, acreage, production costs from National Agricultural Statistics Service; crop futures prices data from the Chicago Board of Trade; ethanol production data from the U.S Department of Energy; and the treasury rates and inflation rate data from the bureau of labor statistics. State-level data across the U.S. from 1980 to 2011 were employed in our research. There are missing observations in our dataset for a number of states. Due to the requirement of a balanced panel, we only chose states whose observations are available for all years. The resulting data includes 42 out of 50 U.S states across 32 years, and 1344 observations are used in our study.

4.1 Variable Descriptions

4.1.1 Farmland Values

The dependent variable in our study is the farmland values per acre. It is the estimated value at which all land used in agricultural production could be sold under current market conditions. All data are estimated as January 1st values. In order to build a uniform timeline for our study, we assume the estimates on farmland values are made at the beginning of every year.

4.1.2 Expected Market Returns

4.1.2.1 Expected Net Farm Incomes

Net farm incomes are measured as the total agricultural net cash incomes subtracted by the government payments. Both the net cash incomes and the government payments data are released from the USDA-Economic Research Service (ERS). The net farm incomes measure all farm-related returns to farm operators after costs have been paid. The expected net farm income is calculated as the arthritic mean of the realized net farm incomes in the five preceding years, which can be formulated as:

$$E_{t-1}NFI_t = \frac{\sum_{k=t-5}^{t-1} NFI_k}{5},$$

where NFI_k represents realized net farm incomes at time *k*, and $E_{t-1}NFI_t$ represents the expected net farm incomes at time *t*, given information available at time *t-1*. In our study the net farm incomes per acre is used as an independent variable and it is calculated as the total net farm incomes in a state divided by the farmland area.

4.1.2.2 Expected Crop Profits

We constructed a crop profits index using various crop production factors including crop yields, prices, production costs and acreage. This index decomposes the expected market returns into separate parts. Farmland values reflect people's projection about how much returns can be generated in the future. Before a parcel of farmland is bought, farmers do not know exactly what the returns will be, but they can forecast the future cash flows by anticipating different production factors. The index represents a rational expectation by economists, farmers, and investors.

The expected profits on a parcel of farmland, $E(\pi)$, can be expressed as a function of crop yields, prices, and costs as:

$$E(\pi) = \sum_j \{E(Y_j) * E(P_j) - C_j\},$$

where $E(Y_j)$ represents the expected yield for planted crop j; $E(P_j)$ represents the expected price of crop j; and C_j represents the production costs. Total profits generated on a parcel of farmland equal the sum of returns from all crop productions on this land.

Due to the difficulty of including all crops produced in the U.S, we built an index by choosing 5 important crops – corn, soybeans, wheat, rice, and cotton. State-level data were used to conduct our analysis. In order to account for different level of production scales in different states, we built an acreage index for all states.

The adjusted framework can be expressed in the formula below; in a given state *i*, the expected profits per acre, $E(\pi_i)$, can be modeled as:

$$E(\pi_i) = \sum_j \{[E(Y_{j,i}) * E(P_{j,i}) - C_{j,i}] * A_{j,i}\}, \text{ where}$$

i = states;

j = corn, soybeans, wheat, rice, cotton;

$E(Y_{j,i})$ = expected yield per acre of crop j in state i;

$E(P_{j,i})$ = expected price of crop j in state i;

$C_{j,i}$ = costs per acre for producing j in state i.

$A_{j,i}$ = acreage index of crop j in state i,

$$= \frac{\text{acres engaged to produce crop j}}{\text{total acres engaged to produce corn, soybeans, wheat, rice, and cotton}}$$

The denominator is not the total agricultural land area to grow all crops in state i, but the total acreage to grow corn, soybeans, wheat, rice, and cotton; because only these five crops are included in our model, the sum of the acreage index should equal 1 in a state, or $\sum_j A_{j,i} = 1$.

Crop Acreage

The crop acreage across 42 U.S. states during 32 years is listed in table 4.1 below.

Crop	**No. Obs**	**Mean**	**St.Dev**	**Min**	**Max**
Corn	1312	1911.802	2974.334	14	14400
Soybeans	928	2332.112	2566.81	7	11000
Wheat	1312	1674.274	2649.519	7	14100
Rice	192	506.2135	416.4106	35	1791
Cotton	512	799.9414	1350.097	0.3	7873

Table 4.1 Statistics of Production Acreage data for Corn, Soybeans, Wheat, Rice, and Cotton.

Note: numbers of observations are different because some states may not grow all five crops.

Detrending Yields in the Crop Profits Index

Due to the strong variation in crop yields over time, a detrending framework was adopted to model expected crop yields. Given a specific parcel of farmland, for crop *j*, the expected yield at time *t* should reflect the historical trend and the crop yields in all years before time *t*. The model can be expressed as follows:

$$E(Y_{j,t}) = \frac{\sum_{m=1975}^{t} \{Y_{j,m} + \Delta_{j,t} * (t-m)\}}{t-1974}; \text{ (t = 1980, 1981, 1982 … 2011)},$$

where $E(Y_{j,t})$ represents the expected crop yield at year t; $Y_{j,m}$ represents the real crop yield at year m; and $\Delta_{j,t}$ represents the yield trend from year 1 to t. Geometrically, Δ corresponds to the slope of the linear regression line of annual crop yields; and statistically it can be interpreted by the formula below:

$$\Delta = \frac{n(\sum T * Y) - (\sum T)(\sum Y)}{n(\sum T^2) - (\sum T)^2}, \text{ where } Y \text{ is yield and } T \text{ is time.}$$

We employ crops yields data from 1975 to 2010 to build our detrending yield framework. The expected yield at time *t* (*t* = 1980, 1981, 1982 … 2011) can be modeled by using the realized-value yields from 1975 to t-1, as indicated in the equations above. Table 4.2 below summarizes the data of expected yields of different crops across 42 states from 1980 to 2011.

Crop	**No. Obs**	**Mean**	**St.Dev**	**Min**	**Max**	**Units**
Corn	1312	91.87505	32.21341	7.497954	172.1288	Bu/Acre
Soybeans	928	30.76815	7.385966	10.82566	50.21594	Bu/Acre
Wheat	1312	42.60571	16.43951	10.756	103.0671	Bu/Acre
Rice	192	58.77664	13.41496	35.3952	87.32186	CWT/Acre
Cotton	512	711.8872	246.767	219.6572	1419.517	Pound/Acre

Table 4.2 Statistics of Production Yields data for Corn, Soybeans, Wheat, Rice, and

Cotton.

Note: numbers of observations are different because some states may not grow all crops.

Expected Crop Prices in the Crop Profits Index

Expected crop prices can be projected using the futures prices. The expected price of crop *j* in state *i* can be modeled as the crop's futures price adjusted by a state specified factor:

$$P_{j,i} = \text{Futures}_j + \lambda_i .$$

λ_i denotes an adjustment factor which explains the price difference between the observed state i and the basis state. In our model the basis state is defined as the place in the futures contract where a given crop is delivered to. The futures price to large extent reflects people's expectation on what price a given crop will be traded at on the settlement date. Meanwhile, crop prices vary across states; in order to capture the across-state price differences, we introduced λ_i which equals the average of historic price differences between state *i* and basis state. Table 4.3 shows the basis states for corn, soybeans, wheat, rice and cotton.

Crop	Basis State
Corn	Illinois
Soybeans	Illinois
Wheat	Illinois
Rice	Arkansas
Cotton	Texas

Table 4.3 Basis States for Corn, Soybeans, Wheat, Rice and Cotton (Source: Futures Contract Specifications, Chicago Board of Trade)

Production Costs in the Crop Profits Index

We use realized-value cost data in our model. While the crop yields and prices are stochastic, depending on a number of agricultural, weather, and economics factors;

farmers are able to know most production costs at the beginning of the year. The crop production costs included:

- Operating costs such as seeds, fertilizer, chemicals, custom operations, utilities, repairs; and
- Overhead costs such as hired labor with opportunity costs of unpaid labor, capital recovery of machinery and equipments, taxes and insurances.

Farmland rents and opportunity costs of land are excluded from total costs, due to their strong correlation with the dependent variable. In addition, interest costs are also omitted in this measure because interest rates exert great impact on farmland values. Thus, we use interest rates as a separate independent variable in the model.

NASS only provides country-level crop costs data. Summary of the crop costs data are listed in the table 4.4 below.

Crop	No. Obs	Mean	St.Dev	Min	Max
Corn	32	254.6444	86.13921	128.33	479.36
Soybeans	32	162.6208	41.70927	111.91	262.33
Wheat	32	132.7153	43.09351	87.5	233.43
Rice	32	494.3256	106.4398	333.05	771.28
Cotton	32	439.9584	103.2796	288.14	672.99

Table 4.4 Statistics of Production Costs data for Corn, Soybeans, Wheat, Rice, and Cotton.

For tractability, production costs per acre for a given crop are assumed to be the same across states. The total crop production costs per acre in a given state are the acreage-weighted average costs of all five selected crops.

4.1.3 Expected Government Payments

We obtain government payment data from the USDA-Economic Research Service (ERS). The data include federal and state payments received by farmers from all direct payments programs. The expected government payments are calculated as the arthritic mean of the realized government payments in five previous years, and can be formulated as:

$$E_{t-1}GP_t = \frac{\sum_{k=t-5}^{t-1} GP_k}{5},$$

where $GP_{i,t}$ represents realized government payments at time *t*, and $E_{t-1}GP_{i,t}$ represents the expected government payments at time *t*, given information available at time *t-1*. In our study the government payments is measured at per-acre level and it is calculated as the total government payments divided by farmland area.

4.1.4 Beale Index

The Beale Rural-Urban Continuum Code is employed as an index to measure the level of urbanization. Based on size, degree of urbanization and proximity to metro areas, all counties in the U.S. are divided into 10 categories. Each county is attached with a Rural-Urban Continuum Code; a code of 0 indicates a county is among the most metropolitan areas; while a code of 9 indicates it is among the most rural ones. The codes have been updated every 9 years since 1974. In our study, for each year we employ the most recent updated codes. The state-level urban index is constructed by taking the arithmetic mean of Beale codes from all counties in a state. Accordingly, states with more urbanized counties such as New Jersey, New York, and California have lower index values; while states in rural areas such as South Dakota and North Dakota have higher index values.

4.1.5 Ethanol Production

Ethanol production data are from the U.S. Department of Energy. The amount of annual ethanol production in the entire U.S, and is measured in millions of gallons.

4.1.6 Real Interest Rates

Real interest rates are measured as the difference between nominal interest rates and inflation rates. 30-year constant maturity treasury (CMT) rates are employed as nominal interest rates because they serve as a measure of long-term borrowing rates. The CMT rates and inflation rates data are from the Bureau of Labor Statistics.

4.1.7 Production Risk

In our study, the unitized risk (also known as variation coefficient) of market returns is used as an indicator of production risk. The calculation for the unitized risk is shown in the equation below:

$$\text{UnitizedRisk} = \frac{\text{Standard Deviation (Market Returns)}}{\text{Mean (Market Returns)}}$$

4.1.8 Farm Bill Periods

A farm bill refers to a federal support policy for farm programs across multiple years. In our study, the farm bill periods are introduced as dummy variables, with FB_1=1981; FB_2=1982-1984; FB_3=1985-1989; FB_4=1990-1995; FB_5=1996-2001; FB_6=2002-2007; FB_7=2008-2011.

4.2 Summary of Variables

Table 4.5 below provides a summary with descriptive statistics for the dependent and independent variables.

Variables	No. Obs	Mean	St. Dev	Min	Max
Farmland Values	1344	1662.676	1718.455	144	15700
Net Farm Incomes	1344	72.7758	77.0836	0.12	535.5193
Crop Profits	1344	31.53817	112.9381	-293.762	846.1235
Government Payments	1344	12.7464	12.3852	0.1696	79.8426
Ethanol	32	2417.01	3187.564	83.074	13297.9
Beale Code (Urban)	1344	5.4397	1.441	0.619	8.1887
Real Interest Rates	32	3.5287	2.2362	-3.3092	7.9171
Production Risk	1344	0.4673	0.134	0.1924	1.1307
Farm Bill Periods			Dummy Variables		

Table 4.5 Definition and Descriptive Statistics for Dependent and Independent Variables.

5. Results

5.1 OLS Estimations

Our study starts with estimating the traditional single-equation capitalization models using ordinary least squares. The estimations regress farmland values against market returns, government payments, risk, real interest rates, urban influencing codes, and ethanol production. Models (1) (2) (3) use net farm incomes as an indicator of market returns and the regression results are presented under columns (1) (2) (3) in Table 5.1. All three models demonstrate positive and highly significant results for the net farm incomes and the government payments variables. However, while the estimates for net farm incomes are extremely consistent in magnitude, the government payments are inconsistent, ranging from 10.3 to 14.2. As indicated in the previous section, this result may be due to the endogeneity issues in the government payments variable, and therefore the instrumental variable approach is needed to further address this issue. Coefficient estimates for other variables including risks, urban influence, interest rate, and ethanol production all demonstrate expected signs in models (1) (2) and (3). All estimates are statistically significant except for the risk variable in model (2).

Model (4) uses crop profits index as an indicator of market returns instead of net farm incomes. The estimation results are shown under column (4) in Table 5.1. As we can observe, the coefficient estimate of crop profits index is insignificant, indicating that the crop profits index we build in this research is not a good indicator of market returns. The reason for this is in some states such as Arizona and Florida, the production of selected crops (corn, cotton, rice, soybeans, and wheat) only takes a small portion of the total crop productions, while the major crops in these states are not accounted for in our index. Therefore the per-acre profits of selected crops cannot represent the weighted averaged profits of all crops produced in those states. Nevertheless, in future studies, a more appropriate crop profits index can be constructed by including a larger number and more representative crops in the “index-constructing pool.” As long as more crops (essentially the major crops of each state) are included in the “pool,” the index will be simulated close to the actual agricultural income and serve as a good indicator of market returns. In our study, however, we will use the net farm incomes as an indicator and all analyses in this chapter are conducted using net farm incomes.

	Original OLS Model	Model with Spatial Expansion Measures	Model with Distance to Midwest Measures	OLS Model with Crop Profits Index
Variables	**Model (1)**	**Model (2)**	**Model (3)**	**Model (4)**
Intercept	2792.95*** (16.76)	2733.507***(7.02)	2810.633*** (17.02)	5053.98*** (29.37)
Net Farm Incomes	10.887*** (24.39)	10.238*** (23.46)	10.911*** (24.66)	
Crop Profits				0.003 (0.01)
Government Payments	14.194*** (6.39)	11.889*** (5.37)	10.266*** (4.4)	27.923*** (10.76)
Risks	-464.625** (-2.39)	-210.841 (-1.12)	-377.953* (-1.95)	294.704 (1.25)
Interest	-47.506*** (-4.2)	-49.188*** (-4.55)	-49.752*** (-4.43)	-44.392*** (-3.26)
Urban (Beale)	-366.349*** (-15.92)	-363.686*** (-16.27)	-368.409*** (-16.15)	-739.232*** (-35.69)
Ethanol	0.113*** (12.76)	-1.693*** (-4.14)	0.169*** (11.95)	0.121*** (11.08)
Ethanol* Coodx		-0.015*** (-3.62)		
Ethanol* Coordy		0.056*** (5.53)		
Ethanol* Coodx* Coordy		0.0005*** (4.72)		
Ethanol* Dist			-0.004*** (-5.05)	
Adjusted R^2	0.7372	0.7605	0.7420	0.7022
Num. Obs	1344	1344	1344	1344

Table 5.1: OLS Estimation Results for Traditional Single-Equation Income Capitalization Models.

Note: Asterisks (*, **, ***) indicate that the statistic is significantly at the confidence level of 10%, 5%, and 1%, respectively. T-statistics are given in parenthesis.

5.2 Estimations Using Instrumental Variable Approach

The instrumental variable approach is then used to conduct the estimation and the results are presented in table 5.2, under column (5) (6) (7), respectively. This approach starts with estimating the government payment equation; predicted values are then used to estimate the farmland value equation. Estimation results of the government payments equation demonstrate significant results for all farm bill period dummy variables, except for *FB2*.

The estimation results from the farmland value equation demonstrate that, both net farm incomes and government payments have highly significant estimated coefficients with positive signs, and all three models present highly consistent results. All other independent variables have highly significant estimates as well. As expected, farmland values are negatively related with real interest rates and production risks; and the negative estimated coefficients of urban influence code indicate that higher urbanization level can result in higher farmland values because a lower urban influence code represents higher urbanization levels.

Estimation from specification (5) demonstrates that ethanol production has a positive impact on farmland values; if ethanol production increases by one million gallons, the farmland value will increase by approximately $0.1. As mentioned above, this model is based on the hypothesis that ethanol production exerts the same level of impact across different states in the U.S. However, this hypothesis may be challenged by the geographic-distribution pattern of ethanol plants. We further explore the geographic influence of ethanol production across the country using models (6) and (7), and the results demonstrate significant coefficient estimates of all geographic-related variables (i.e., $Ethanol*Coordx$, $Ethanol*Coordy$,and $Ethanol*Coordx*Coordy$ in model 6, and $Ethanol*Dist$ in model 7), which proved the necessity of considering geographic effects while measuring the impact of ethanol production. To explain the relationship between farmland values and ethanol production, we need to measure the joint impact of all ethanol-related variables. For a specific state *i*, the joint impact of ethanol productions can be expressed as:

Model (6): $$\beta_{joint,i} = \beta_{eth} + \beta_{eth*x}Coordx_i + \beta_{eth*y}Coordy_i + \beta_{eth*x*y}Coordx_i*Coordy_i,$$

And Model (7): $$\beta_{joint,i} = \beta_{eth} + \beta_{eth*dist}Dist_i,$$

where $\beta_{joint,i}$ represents the joint coefficient of ethanol production at state *i*; β_{eth}, β_{eth*x}, β_{eth*y}, and $\beta_{eth*dist}$ represent the estimated coefficients corresponding to $Ethanol$, $Ethanol*Coordx$, $Ethanol*Coordy$, $Ethanol*Coordx*Coordy$, and $Ethanol*Dist$, respectively.

Figure 5.1 and 5.2 below map out the estimated joint coefficients of ethanol production for different states in the U.S. Both figures demonstrate obvious spatial effects with ethanol production. However, the estimation results differ in two aspects: 1)

Estimated coefficients from the two models vary in magnitude -- in model (5) $\beta_{joint,i}$ ranges from -0.03 to 0.27; whereas in model (6) $\beta_{joint,i}$ ranges from 0.01 to 0.16; 2) As presented in the figures below, the estimated spatial effects with ethanol impacts follow different patterns – model (6) demonstrates a declining trend from northeast to southwest; but model (7) demonstrates a radial dispersal pattern with the Midwest as a central.

Our study is in favor of model (7) because it better reflects the real world situations. According to data released from USDA, more than 80% of the ethanol plants are located in the Midwest area. Large ethanol production scale would boost crop (especially corn) demand and productions and therefore increase the farmland value in the nearby areas. Estimation results from model (7) explain this effect in that it shows the further away a state is located from the Midwest, the less impact ethanol productions will have on farmland values. Even though model (6) has slightly higher explanatory power than (7), it has to follow a straight longitude-latitude pattern due to the limitation of model specification, and does not explain the real situation well. The higher R-squared may because it absorbs some of the spatial dependence in the dependent variable (farmland values) and also be the reason why it demonstrates higher estimated coefficients $\beta_{joint,i}$ than (7). Therefore, models with distance to the Midwest measures are considered as a better specification to measure farmland values.

	Model without Spatial-measuring Independent	**Model with Spatial Expansion Measures**	**Model with Distance to Midwest Measures**

Variables	Variables		
Variables	Model (5)	Model (6)	Model (7)
Farmland Value Equation			
Intercept	2866.794***	2774.333***	2840.121***
	(17.21)	(17.36)	(17.33)
Net Farm Incomes	10.431***	9.49***	10.25***
	(21.12)	(19.6)	(21.07)
Gov. Payments	26.439***	26.496***	25.474***
	(5.02)	(5.29)	(4.91)
Risk	-901.158***	-600.201***	-708.712***
	(-4.55)	(-3.16)	(-3.6)
Interest Rate	-44.397***	-44.832***	-44.593***
	(-3.85)	(-4.09)	(-3.93)
Urban (Beale)	-363.168***	-360.035***	-369.754***
	(-15.7)	(-16.14)	(-16.24)
Ethanol	0.105***	-1.388***	0.17***
	(10.44)	(-3.43)	(12.33)
Ethanol* Coodx		-0.012***	
		(-2.94)	
Ethanol* Coordy		0.05***	
		(4.93)	
Ethanol*Coodx* Coordy		0.0004***	
		(4.19)	
Ethanol* Dist			-0.005***
			(-6.78)
Adjusted R^2	0.7342	0.7603	0.7429
Num. Obs	1344	1344	1344
Government Payment Equation			
Intercept	-5.628**	-5.628**	-5.628**
	(-2.49)	(-2.49)	(-2.49)
Net Farm Incomes	0.045***	0.045***	0.045***
	(9.08)	(9.08)	(9.08)
Urban (Beale)	0.898***	0.898***	0.898***
	(3.49)	(3.49)	(3.49)
Ethanol	-0.0003	-0.0003	-0.0003
	(-1.21)	(-1.21)	(-1.21)

FB_2=1	0.673	0.673	0.673
	(0.39)	(0.39)	(0.39)
FB_3=1	6.583***	6.583***	6.583***
	(3.88)	(3.88)	(3.88)
FB_4=1	9.472***	9.472***	9.472***
	(5.61)	(5.61)	(5.61)
FB_5=1	10.237***	10.237***	10.237***
	(6.03)	(6.03)	(6.03)
FB_6=1	21.514***	21.514***	21.514***
	(11.06)	(11.06)	(11.06)
FB_7=1	20.653***	20.653***	20.653***
	(6.24)	(6.24)	(6.24)
Adjusted R^2	0.3492	0.3492	0.3492
Num. Obs	1344	1344	1344

Table 5.2: Estimation Results for Instrumental Variable Approach.
Note: Asterisks (*, **, ***) indicate that the statistic is significantly at the confidence level of 10%, 5%, and 1%, respectively. *T*-statistics are given in parentheses.

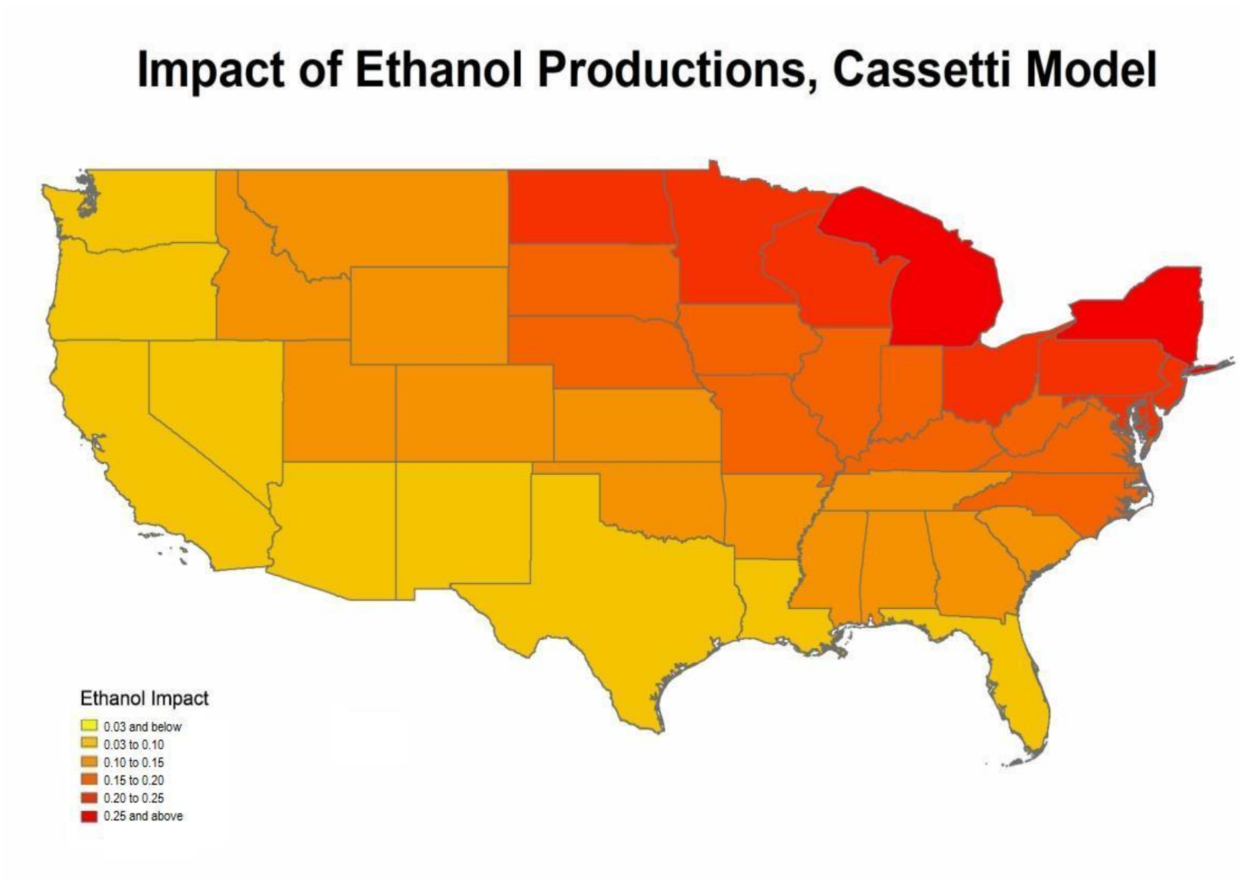

Figure 5.1 Impact of Ethanol Productions, Model (6)

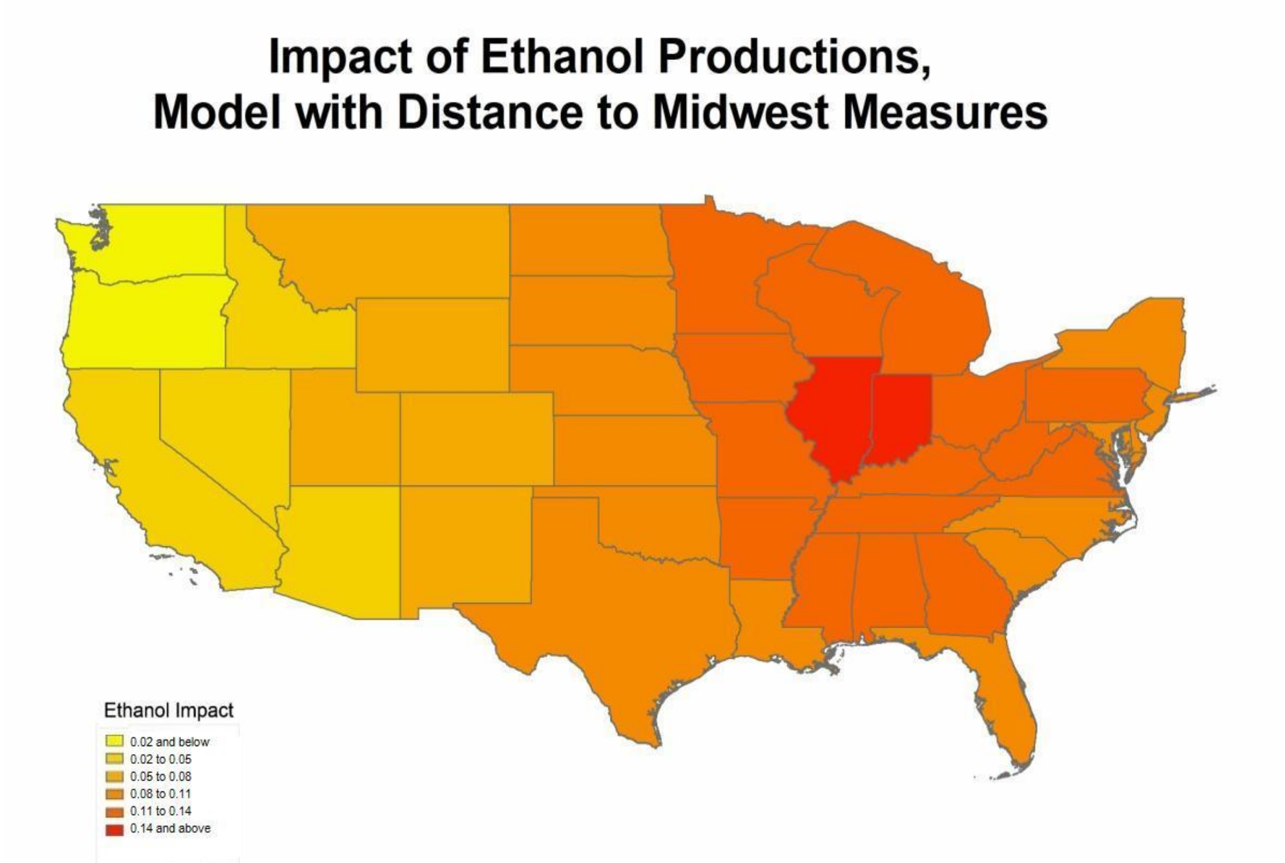

Figure 5.2 Impact of Ethanol Productions, Model (7)

Following Freeman (1993), an underlying assumption of this income capitalization model is that all regions analyzed in the study are independent markets and do not exert any influence on each other. More specifically in our farmland values case, the assumption indicates that the farmland values of a given state can be exclusively explained by the attributes in that state, not being affected by factors in other regions. Violation of the assumption or ignoring the spatial dependence may result in a biased estimation. However, Huang et al. (2006) and Hardie et al. (2001) indicate in their studies that the land values tend to be spatially correlated. Our study conducted tests on spatial dependence using Moran’s I statistics; table 5.3 shows the existence of spatial correlation in farmland values across states in the U.S, during the period of 1980 through 2011.

Year	Moran's I Statistics	z-Score	Probability
1980	0.343450	3.8162719	0.0001355
1981	0.319996	3.5761962	0.0003487
1982	0.281082	3.1778648	0.0014838
1983	0.296197	3.3325781	0.0008606
1984	0.304123	3.4137119	0.0006409
1985	0.285482	3.2228994	0.0012691
1986	0.351664	3.9003523	0.0000961
1987	0.377203	4.1617787	0.0000316
1988	0.365200	4.0389122	0.0000537
1989	0.360235	3.9880955	0.0000666
1990	0.334022	3.7197739	0.0001995
1991	0.280881	3.1757995	0.0014944
1992	0.287885	3.2960549	0.0009806
1993	0.267359	3.0556100	0.0022462
1994	0.267385	3.0558788	0.0022442
1995	0.277084	3.1554771	0.0016025
1996	0.275900	3.1433144	0.0016706
1997	0.271660	3.0997776	0.0019368
1998	0.281707	3.2029541	0.0013604
1999	0.297674	3.3669298	0.0007602
2000	0.314142	3.5360485	0.0004062
2001	0.319906	3.5952379	0.0003242
2002	0.329118	3.6898358	0.0002245
2003	0.335915	3.7596433	0.0001702
2004	0.382405	4.2370683	0.0000227
2005	0.439230	4.8549852	0.0000012
2006	0.440900	4.8721809	0.0000011
2007	0.431801	4.7784622	0.0000018

2008	0.426074	4.7194721	0.0000024
2009	0.429517	4.7549398	0.0000020
2010	0.410706	4.5611699	0.0000051
2011	0.408156	4.5349060	0.0000058

Table 5.3 Moran's I Statistics on Farmland Values from 1980 to 2011.

Results from Moran's I statistics demonstrate strong spatial autocorrelation on farmland values across states, with $p<0.01$ in all years. The statistics are relatively stable in the 32 years of our study, ranging from 0.26 to 0.44, indicating consistent spatial autocorrelation in different years. Because ignoring spatial dependence in the empirical model may impact the consistency and efficiency of the estimation results (Kim et al., 2003); it is necessary to use spatial lag and spatial error specifications in the study on farmland values.

5.3 Spatial Lag and Spatial Error Specifications

Estimation results using spatial lag and spatial error specifications with random and fixed effects are shown in table 5.4 below, under column (8) through (11), respectively. Similar to models (5) through (7) discussed above, to avoid the effects of endogeneity, an equation of government payments was estimated initially, and the predicted values from that equation are used as in place of government payments in following spatial models. Estimation results of the government payments equation were the same as in models (5) through (7) presented in table.

	Spatial Lag (FE)	**Spatial Error (FE)**	**Spatial Lag (RE)**	**Spatial Error (RE)**
Variable	**Model (8)**	**Model (9)**	**Model (10)**	**Model (11)**

Intercept	2956.202*** (13.77)	4681.761*** (17.71)	2880.789*** (3.84)	4641.051*** (5.75)
Net Farm Incomes	11.718*** (20.2)	14.134*** (17.7)	23.165*** (19.56)	23.133*** (15.25)
Government Payments	24.022*** (3.88)	34.417*** (3.18)	31.042*** (3.67)	84.145*** (5.04)
Risk	-919.759*** (-3.93)	-576.011** (-1.93)	-4342.225*** (-9.16)	-3793.076*** (-6.7)
Interest Rate	-44.223*** (-3.28)	-68.326*** (-2.84)	-38.517** (-2.49)	-94.954*** (-2.65)
Urban (Beale)	-386.014*** (-12.8)	-653.763*** (-18.15)	-236.287** (-2.12)	-420.796*** (-3.42)
Ethanol	0.157*** (8.68)	0.255*** (9.02)	0.113*** (4.91)	0.314*** (6.94)
Ethanol* Dist	-0.004*** (-4.7)	-0.007*** (-5.14)	-0.002* (-1.77)	-0.005** (-2.08)
ρ	0.172*** (6.08)	0.33*** (9.68)	0.532*** (21.97)	0.545*** (19.7)
Adjusted R^2	0.7411	0.7422	0.7146	0.7256
Num. Obs	1344	1344	1344	1344

Table 5.4 Estimation Results for Spatial Lag and Spatial Error Models.
Note: Asterisks (*, **, ***) indicate that the statistic is significantly at the confidence level of 10%, 5%, and 1%, respectively. T-statistics are given in parenthesis.

FE in parenthesis denotes fixed effects estimates; RE in parenthesis denotes random effects estimates.

The spatial autoregressive coefficients (ρ) demonstrate positive and highly significant results in all models. There is strong evidence of spatial autocorrelation both in the dependent variable and in the error terms. The models are under the assumption that the spatial effects remain constant over time (Woodard et al, 2010).

All four spatial models have reasonable explanatory power, with adjusted R^2 ranging from 0.71 to 0.74 approximately. Estimated coefficients from all models are statistically significant and have signs as expected, same as in models (5) through (7). However, it is necessary to mention that the coefficient estimates from the random effect models vary significantly in magnitude versus the estimates from the fixed effect models, therefore proper tests need to be conducted in order to select the best estimation model.

5.3.1 Wu-Hausman Tests

To choose the most proper specification between the random and fixed effects, we conduct a Wu-Hausman test, the statistics of which can be expressed as,

$$Hausman = [\widehat{\beta}_{FE} - \widehat{\beta}_{RE}]'[Var(\widehat{\beta}_{FE}) - Var(\widehat{\beta}_{RE})]^{-1}[\widehat{\beta}_{FE} - \widehat{\beta}_{RE}]$$

The Wu-Hausman statistics follows a chi-squared distribution with degrees of freedom equal to the rank of the Var($\hat{\beta}_{FE}$) and Var($\hat{\beta}_{RE}$) matrix (Hausman, 1978). As Hausman indicated, this is a useful test to differentiate between the random effects and fixed effect models. The null hypothesis of the test suggests that both the random effects and fixed effects models are consistent; if the null hypothesis is true, the random effects models would be preferred because of higher efficiency. Alternatively, if the null hypothesis is rejected, the fixed effects models would be more appropriate because they are consistent; while the random effects models are not.

Results of Wu-Hausman tests are shown in table 5.5. The *p*-values of 0.000 indicate we should reject the null hypothesis in both tests, which means the random effects models are inconsistent in the estimations. This can be also the reason why estimations with random effects yield different results than estimations with fixed effects. Therefore, fixed effects models are preferred in our study.

Wu-Hausman Tests	**Wu-Hausman Statistics**	**Degrees of Freedom**	**Probability**
Spatial Lag	587.1527	9	0.000
Spatial Error	91.9612	9	0.000

Table 5.5 Wu-Hausman Test on Random Effects and Fixed Effects Specifications.

Note: H_0: Both random effects and fixed effects estimates are consistent;
H_1: Random effect estimates are not consistent.

5.3.2 Lagrange Multiplier (LM) Tests and Robust LM Tests

To choose an appropriate framework between the spatial error model and the spatial lag model, we conducted two Lagrange Multiplier (LM) tests and two robust LM tests. The (Robust) LM tests follow a Chi-squared distribution with degree of freedom of 1 (Anselin, 1988); as Anselin indicated, results of these Lagrange Multiplier tests can serve as a good guild to decide which model, between spatial lag and spatial error, is preferred in our study.

Table 5.6 present the results of the Lagrange Multiplier tests and respective robust LM tests. The P-values of 0.000 from two LM tests (Line 1 and Line 2) indicate that we should reject the null-hypothesis of no spatial dependence in the original OLS model, meaning the existence of spatial autocorrelation in both the error terms and the dependent variable.

Two additional robust LM tests (Line 3 and Line 4) were conducted. As shown in the table 5.6, only the robust lag test is insignificant, indicating with the presence of the spatial error term, the spatial dependence in the dependent variable is highly likely to disappear. Therefore the spatial error model is preferred in our study.

(Robust) LM Tests H_0: No Spatial Dependence	z-Value	Probability
(1) Lagrange Multiplier, Lag	54.8798	0.000
(2) Lagrange Multiplier, Error	109.1643	0.000
(3) Robust Lagrange Multiplier, Lag	0.1730	0.677
(4) Robust Lagrange Multiplier, Error	54.4575	0.000

Table 5.6 Lagrange Multiplier (LM) Tests and Robust LM Tests on Spatial Autocorrelation
Note: H_0: No spatial correlation in the estimation; H_1: Presence of spatial correlation in the estimation.

5.3.3 Estimation Results

As discussed above, the (robust) Lagrange Multiplier tests and Wu-Hausman tests demonstrate that, the spatial error model with fixed effects (model 9) is in favor of other spatial models used in our study. Estimation result is shown in table 5.4 under column (9). Coefficient estimates for all independent variables are highly significant with signs as expected. The farmland values are positively related with expected crop profits and government payments, with one dollar increase in expected crop profits and government payments, the farmland values will increase by $14.13 and $34.42, respectively.

Additionally, estimation results confirmed our expectation that, the real interest rate exerts negative impact on farmland values; with one percentage increase in real interest rates, the farmland values will decrease by approximately $68.33. The Beale index as a measure of urban influences is found to have negative relationship with farmland values. Even though not directly related with farm incomes or agricultural productions, the urbanization characteristics are highly influential in that the farmland values may increase due to the potential residential or entertainment conversions, or increase in population densities. The results demonstrated that the farmland values are higher in more urbanized areas.

The coefficient estimates on ethanol production variables presented similar results as in the non-spatial models. The negative sign of coefficient on $Ethanol * Dist$ indicates that the impact of ethanol productions on farmland values is stronger in states around the Midwest area. Figure 5.5 below presents the ethanol impact in different states across the U.S., depicting the joint estimated coefficients of $Ethanol$ and $Ethanol * Dist$. With one million gallons increase in ethanol production, the farmland values will increase by \$0.02 to \$0.24, depending on different states.

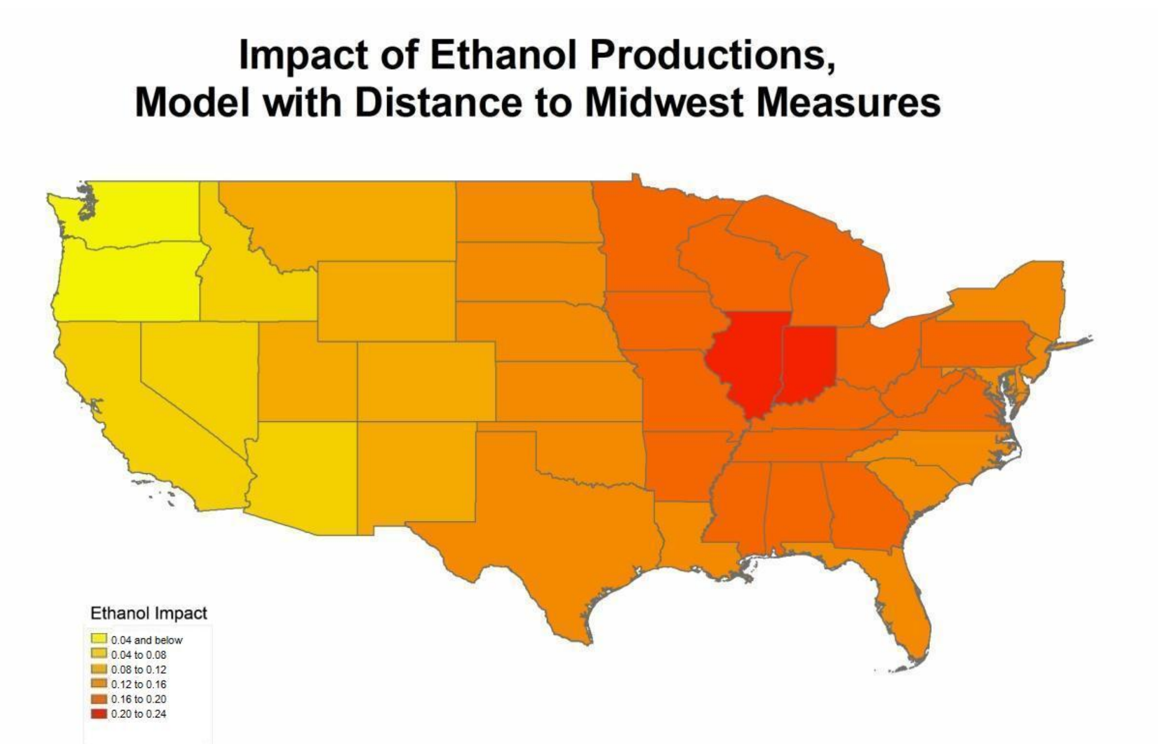

Figure 5.5 Ethanol Impacts across States

6. Proposal on Future Research and Farmland Values Forecasts

In the previous chapter, we have determined an appropriate approach in estimating the U.S. farmland values, using the fixed-effect spatial error econometric model. As discussed above, in future studies, this approach can be further improved by constructing an appropriate crop profits index and using that index as an indicator of market returns to estimate the econometric model. Because the profits index can be broken down into various crop production factors, estimation using crop profits index directly links farmland values with the factors, such as crop yields, prices, and production costs, which can provide us with a deeper insight of how the underlying determinants may affect farmland values. In addition, once an appropriate index can be constructed, we will be able to forecast the future farmland values according to the changes in determinants, such as crop prices and interest rates. The basic theories and methodologies for conducting the forecasts are described below.

Following Bessa et al. (2012), the farmland values forecast problem consists in determining the probability density function (PDF) of farmland values at time *t* for each look-ahead time *t+k*. Having the forecasted joint probability density function of the explanatory variables, the forecast function of farmland values can be formulated as:

$$f_L(L_{t+k} \mid X = x_{t+k|t}) = \frac{f_{L,X}(L_{t+k}, x_{t+K|t})}{f_X(x_{t+k|t})},$$

where L_{t+k} is the forecasted farmland values at time *t+k*, $x_{t+k|t}$ are a vector of forecasted explanatory variables at time *t+k* simulated at time *t*, $f_{L,X}$ is the joint density function of the farmland values and explanatory variables, f_X is the forecasted joint density function of the set of explanatory variables, and $f_L(L_{t+k} \mid X = x_{t+k|t})$ is the density function of farmland values for look-ahead time *t*+*k*, which is to be forecasted.

In the real practice of the forecasting process, the look-ahead joint probability density function for explanatory variables can be estimated initially, using data available in/before year *t*. Then the forecasted values for the explanatory variables will be fitted into the farmland values forecast model to build a density function of farmland values.

For tractability, one may conduct the forecast mainly only focusing on the changes in two sets of explanatory factors: crop prices and interest rates. These variables vary significantly over time and are considered to have made greatest contributions to the changes in farmland values during recent years (Gloy, 2012). Other independent variables in the model can be assumed to remain constant, given they do not impact the volatilities to large extent.

Crop futures prices and options prices at time *t* can be used to simulate the joint distribution for expected crop prices. By introducing the options premiums and inverting the Black-Scholes model (for details see Black and Sholes, 1973), one will be able to calculate the volatilities for expected crop prices. The volatilities combined with current futures prices provide us with the distribution of each crop. Log-normal distributions can be assumed for crop prices.

Similar approach can be used to simulate the distribution of interest rates, using 30 year treasury bond futures and options prices. It is necessary to mention that the treasury bond futures are quoted in terms of points (each point represents $1000), instead of yield. A log-normal distribution is assumed for treasury bond prices. The Black-Scholes model can be used to invert the treasury bond options premiums to calibrate volatilities. Having the volatilities and the futures prices, one is able to simulate the distribution for treasury bond prices (in terms of dollars), and to convert the prices back to bond yield, which can be used as the indicator of interest rates.

Moreover, historical crop prices and treasury bond yields can be used to measure the correlations. The Guassian copula is a good approach to account for the correlations and simulate the joint distribution for expected crop prices and interest rates (for details about the Gaussian copula see Nelsen, 2006). In addition, one should also simulate the state-specific distributions for error terms from the farmland value estimation model. This need to be done because the error terms are a primary source of volatility in farmland value forecast; ignoring the error terms may underestimate the variance of the forecasted distributions.

With the simulated joint density distributions for expected crop prices and interest rates, as well as the simulated distributions for error terms, the state-level farmland values can be forecasted. The equation used for the forecast is shown as follows:

$$L_t = \beta^{Est} X_t + (I_{NT} - \rho^{Est} W_{NT})^{-1} * u_t,$$

where L_t is a vector of forecasted farmland values across states at time t; I_{NT} is an identity matrix; W_{NT} is a weight matrix; ρ^{Est} is the estimated spatial autoregressive coefficient; β^{Est} is a vector of estimated coefficients for explanatory variables; X_t is a matrix of simulated values for the explanatory variables; and u_t is a vector of bootstrapped errors.

Nevertheless, it is necessary to mention that in our forecasting framework, all other independent variables except crop prices and interest rates are assumed to remain constant in the period of the forecast time-horizon. In reality, the variables are likely to change over time and may to some extent exert impacts on future farmland values. Unfortunately, some explanatory variables have high level of uncertainty and are relatively difficult to forecast. For example, ethanol production is determined by various economic factors including international demand, technology, and policy, therefore forecasting ethanol production requires large efforts. However, to provide a more accurate forecast on farmland values, a better forecasting framework should be developed and all explanatory variables should be considered and controlled for in the forecasting model.

7. Conclusion

Our research focuses on the study of farmland values employing state-level data in the U.S. during the period of 1980 to 2011. The specific objectives were stated as follows:

1) To explore the most appropriate econometric model on farmland valuation.
2) To determine factors influencing farmland values and to examine the relationships between underlying determinants and farmland values.
3) To test spatial dependencies in farmland values across different states in the U.S.

4) To offer suggestions for future researches and to provide a proposal in forecasting future farmland values according to the changes in the determining factors.

While spatial relationships across state lines lead to important economic questions, the key advances addressed in this thesis are in the novelty of the spatial econometric approaches utilized, the scope and scale of the model implementation, our approach for addressing endogeneity and other estimation issues. Thus, while the economic problem is centered on spatial relationships in farmland pricing, the research problem in our context is centered on the appropriate method to capture these impacts, and in doing so lead to some very informative market driven results about the magnitude of risk in farmland values today compared to earlier periods. Indeed, while a variety of authors such as Gloy have commented on the fundamentals of and risk to current farmland values, our study is the first to employ a scientific, cohesive, and defensible approach to putting a figure to such risks. Consequently, we explored the most appropriate econometric specification for modeling large scale U.S. farmland valuations, and multiple specifications were estimated and tested in the course of the analysis. Our results demonstrate that our spatial instrumental variable approach should be applied when endogeneity issues related to government payments variable and other expectations exist, and when geographically-differentiated impacts from ethanol production (e.g. the haves versus have-not states in relation to measured distances to the Midwest) need to be addressed. Due to the existence of spatial dependencies, we conducted analyses using spatial lag and spatial error models with fixed effects and random effects. A number of tests including

Moran's I tests, (robust) Lagrange Multipliers tests, and Wu-Hausman tests are employed to investigate in the advantages and disadvantages of different spatial models. Results demonstrate that the spatial error model with fixed effects is the most appropriate specification in our study.

Next, we examined the relationships between underlying determinants and farmland values. Estimation using fixed-effect spatial error model yields significant results for all independent variables being studied. The results demonstrate that crop profits and government payments are two sources of farm income that exert positive impact on farmland values, and that a one dollar increase in expected annual crop profits and government payments leads to an increase in estimated farmland values by $14.13 and $34.42, respectively. In addition, farmland values are positively related with urbanization pressures and ethanol production, and are negatively related with interest rates and risk, which are all as expected. Furthermore, the spatial modeling framework allows us to test the spatial dependencies in farmland values across different states. The estimation and test results demonstrate that the farmland values in a state are highly likely to be influenced by neighboring states.

Last, we offered some suggestions for future researches and provided a proposal for forecasting farmland values according to the changes of underlying determinants. We concluded that to further improve the farmland values estimation, an appropriate crop profits index can be constructed and used as an indicator of market returns to estimate the econometric model, because estimation using crop profits index directly links farmland values with the production factors, such as crop yields, prices, and production costs, which can provide us with a deeper insight of how the underlying determinants may affect farmland values in the U.S. In addition, using the crop profits index to estimate the model can facilitate us to forecast the future farmland values according to the changes in determinant variables, such as crop prices and interest rates. The basic theories and methodologies for conducting the farmland value forecasts are described in the thesis, to provide an instruction for future studies.

REFERENCES

Anselin, L. 1988. Spatial Econometrics: Methods and Models. Dordrecht, The Netherlands: Kluwer Academic Publishers, Inc.

Barnett, Barry J. 2010. “The U.S. Farm Financial Crisis of the 1980s.” Agricultural History 74(2): 366-380.

Bessa, Ricardo J., Miranda V., Botterud A., Zhou Z., and Wang J. 2012. “Time-adaptive quantile-copula for wind power probabilistic forecasting.” Renewable Energy 40(1): 29-39.

Black, F., and M. Scholes. 1973. “The Pricing of Options and Corporate Liabilities.” Journal of Political Economy 81(3): 637–654.

Cassetti, E. 1972. “Generating Models by the Expansion Method: Applications to Geographical Research.” Geographical Analysis 4(1): 81–91.

Clark, J. S., K. K. Klein, and S. J. Thompson. 1993. “Are Subsidies Capitalized into Land Values? Some Time Series Evidence from Saskatchewan.” Canadian Journal of Agricultural Economics 41(2): 155-168.

Craig, L. A., R. B. Palmquist, and T. Weiss. 1998. “Transportation Improvements and Land Values in the Antebellum United States: A Hedonic Approach.” Journal of Real Estate Finance and Economics 16(2): 173-189.

Elhorst, J. P., Lacombe, D. J., Piras, G. 2012. “On Model Specification and Parameter Space Definitions in Higher Order Spatial Econometrics Models.” Regional Science & Urban Economics 42(1-2): 211-220.

Gloy, B. 2012. “When Do Farm Booms become Bubbles?” Unpublished paper presented at 2012 Agricultural Symposium, Kansas City.

Gloy, Brent, Michael D. Boehlje, Craig L. Dobbins, Christopher Hurt, and Timothy G. Baker. 2011. “Are Economic Fundamentals Driving Farmland Values?” Choices, the Magazine of Food, Farm and Resource Issues, Agricultural and Applied Economics Association 26(2).

Goodwin, B. K., and F. Ortalo-Magne. 1992. “The Capitalization of Wheat Subsidies into Agricultural Land Values.” Canadian Journal of Agricultural Economics 40(1): 37-54.

Goodwin, B. K., A. K. Mishra, and F. N. Ortalo-Magné 2003. “What's Wrong with Our Models of Agricultural Land Values?” American Journal of Agricultural Economics 85(3): 744-752.

Hardie, I. W., T. A. Narayan, and B. L. Gardner. 2001. “The Joint Influence of Agricultural and Non-farm Factors on Real Estate Values: An Application to the Mid-Atlantic Region.” American Journal of Agricultural Economics 83(1): 120-132.

Hamilton, J. D., and C. H. Whiteman. 1985. “On Testing for Self-Fulfilling Speculative Bubbles.” Journal of Monetary Economics 16(2): 333-373.

Hausman, J. A. 1978. "Specification Tests in Econometrics". Econometrica 46 (6): 1251–1271.

Henderson, J., B.A. Gloy, and M.D. Boehlje. 2011. “Agriculture’s Boom-Bust Cycles: Is This Time Different.” The Economic Review, Kansas City Federal Reserve Bank 96(4): 83-101.

Henderson, J. 2008. “Will farmland values keep booming?” Economic Review -- Federal Reserve Bank of Kansas City 93(2): 81-104.

Henderson, J. and B. Gloy. 2008. “The Impact of Ethanol Plants on Cropland Values in the Great Plains.” Agricultural Finance Review 69(1): 36-48.

Huang, H., G. Miller, B. Sherrick, and M. Gomez. 2006. “Factors Influencing Illinois Farmland Values.” American Journal of Agricultural Economics 88(2): 458-470.

Kapoor, M., Kelejian, H. H., and Prucha, I. R. 2007. “Panel data model with spatial correlated error components.” Journal of Econometrics 140(1): 97-130.

Kim, C. W., T. T. Phipps, and L. Anselin. 2003. “Measuring the Benefits of Air Quality Improvement: A Spatial Hedonic Approach.” Journal Environmental Economics and Management 45(1): 24-39.

Kirwan, B. E. 2009. “The Incidence of US Agricultural Subsidies on Farmland Rental Rates.” Journal of Political Economy 117(1): 138-164

Kropp, Jaclyn D., and Peckham, Janet G. 2012. “Impact of U.S. Agricultural and Ethanol Policies on Farmland Values and Rental Rates.” Unpublished paper presented at Agricultural & Applied Economics Association’s 2012 AAEA Annual Meeting, Seattle, Washington.

Nelsen, R. B. 2006. An Introduction to Copulas, Second Edition. New York, NY, USA: Springer Science & Business Media Inc.

Palmquist, Raymond B. 1989. "Land as a Differentiated Factor of Production: A Hedonic Model and its Implications for Welfare Measurement." Land Economics 65(1): 23-28.

Palmquist, R.B., and L.E. Danielson. 1989. "A Hedonic Study of the Effects of Erosion Control and Drainage on Farmland Values." American Journal of Agricultural Economics 71(1): 55-62.

Phipps, T. 1984. "Land Prices and Farm-Based Returns." American Journal of Agricultural Economics 66(4): 422-429.

Shaik, S., G. A. Helmers, and J. A. Atwood. 2005. "The Evolution of Farm Programs and their Contribution to Agricultural Land Values." American Journal of Agricultural Economics 87(5): 1190–1197.

U.S Department of Agriculture, Economic Research Service (USDA-ERS). Historical U.S. and State-Level Farm Income and Wealth Statistics. Retrived from http://www.ers.usda.gov/data-products/farm-income-and-wealth-statistics/historical-us-and-state-level-farm-income-and-wealth-statistics.aspx#.Un-v33BHIyV

U.S Department of Agriculture, National Agricultural Statistics Service (USDA-NASS). 2012. "Crop Production 2012 Summary."

U.S Department of Agriculture, National Agricultural Statistics Service (USDA-NASS). 2012. "Land Values 2012 Summary."

U.S Department of Agriculture, National Agricultural Statistics Service (USDA-NASS). Agricultural Statistics Annual. Retrived from http://www.nass.usda.gov/Publications/Ag_Statistics/index.asp

Vyn, J. Richard, Zahoor UI Haq, Jeevika Weerahewa, and Karl D. Meilke. 2012. "The Influence of Market Returns and Government Payments on Canadian Farmland Values." Journal of Agricultural and Resource Economics 37(2): 199-212.

Weersink, A., S. Clark, C. G. Turvey, and R. Sarker. 1999. "The Effect of Agricultural Policy on Farmland Values." Land Economics 75(3): 425–439.

Woodard, S., N. D. Paulson, K. Baylis, and J. D. Woodard. 2010. "Spatial Analysis of Illinois Agricultural Cash Rents." Working paper.

Xu, Feng, Mittelhammer, Ron C., Barkley, Paul W. 1993. "Measuring the Contributions of Site Characteristics to the Value of Agricultural Land**.**" Land Economics 69(4): 356 – 369.