Journal of Agribusiness 36, 1 (Spring 2018)

The Impact of a Sorghum-Based Ethanol Plant on Local Sorghum Basis and Cotton Acreage: A Spatial Approach

Bing Liu, Darren Hudson, Jamie Bologna Pavlik, and Michael Farmer

This study examines the impacts of a sorghum-based ethanol plant established in a major cotton-producing area on local sorghum basis and cotton acreage distribution using county-level panel data from 2002 to 2014. Spatial econometric models are employed to account for any spatial dependence. Our results support the conclusion that sorghum basis and cotton acreage within a county depends on characteristics of its neighbors. Specifically, the findings indicate that the sorghum basis increased by 1.5 cents per bushel in the short run resulting from hosting a 40-million-gallon ethanol plant, and a short-run increase by 0.2% in cotton acres over all counties.

Key words: Cotton acreage, Ethanol plant location, Ethanol production, Sorghum basis, Spatial models

Over the past decade, construction of starch-based ethanol plants have expanded rapidly across the United States. From 2007 to 2017, ethanol production capacity in the United States increased from 6.5 to 15.3 billion gallons. As of May 2017, the Renewable Fuels Association (RFA) has listed 213 operational ethanol plants nationwide (RFA, 2017). Changes in U.S. energy policy dramatically boosted both the demand for and supply of ethanol. On the demand side, the Renewable Fuel Standard (RFS) mandates expanded drastically in 2007, calling for an annual consumption of 36 billion gallons of renewable fuels by 2022. The ban of methyl tertiary butyl ether (MTBE), a close substitute of ethanol, as a fuel oxygenate has contributed to the recent growth in demand.[1] As for supply, the U.S. government has provided tax and financing incentives for investments in new ethanol processing plants.

The increasing criticism of the sustainability of first-generation ethanol has increased interests in developing second-generation ethanol from renewable agricultural biomass rather than food-based resources. Among biomass crop plants, sorghum has emerged as a potential feedstock candidate for ethanol production because it meets the energy

[1] MTBE, used as a fuel additive to reduce carbon monoxide emissions, was found to contaminate both ground and surface water and was banned in 19 states by 2004.

Bing Liu is research scientist, Darren Hudson is a professor, Jamie Bologna Pavlik is an assistant professor, and Michael Farmer is an associate professor, all in the Department of Agricultural and Applied Economics, Texas Tech University, Lubbock. We are appreciative of the anonymous reviewer's comments. Any remaining errors are the sole responsibility of the authors. The authors gratefully acknowledge partial funding by Cotton, Inc. and the Larry Combest Endowed Chair in Agricultural Competitiveness, which supported this research.

requirements to be environmentally sustainable, easily adopted by producers, and takes advantage of existing agricultural infrastructure. Currently, roughly 3% of the feedstock required for commercial ethanol production comes from sorghum, with 95% coming from corn (RFA, 2017). Researchers and ethanol producers have shown that sorghum as a feedstock could make a larger contribution to the nation's fuel ethanol requirements (Rooney et al., 2007; Balat and Balat, 2009; Linton et al., 2011).

It is important to note that ethanol has generally not been profitable without the help of subsidies (Hill et al., 2006) so expansion of ethanol production necessarily diverts resources from their highest and best use in a free market and results in substantial deadweight losses (Gardner, 2007). While the producers of sorghum may be better off, it is unclear how the rest of the economy will fair. Therefore, the policy-induced effects are impacting the crop allocation in ways that potentially have adverse impacts in local areas where feedstocks are produced. We, however, are not addressing the broader general equilibrium effects here. Rather, the goal of this paper is to examine the impacts of a sorghum-based ethanol plant in the Texas High Plains, a major cotton-producing area, on local sorghum basis and cotton acreage distribution.

Most of the recent developments in ethanol production have focused on rural areas. Locating an ethanol plant in a small, rural city is expected to potentially benefit the local economy in terms of increased job opportunities, enhanced farmer income through purchases of local farm production to be used as a feedstock, and improved community infrastructure for future potential growth. In places like Iowa, increased demand for corn for ethanol may have only minimal effects on farmers' cropping decisions as they largely grow corn already. However, establishing a sorghum-based ethanol plant in an area with intensive cotton production, like the Texas High Plains, raises important economic issues. Questions concerning how the sorghum supply changes in a region where a new ethanol plant is introduced, how the increased demand for sorghum affects price received for the farmers located close to the ethanol plant, and how the changes in the price affect farmers' planting decisions have increased in importance as the biofuel industry continues to expand.

Diverting sorghum away from its traditional feed and export uses to the production of ethanol affects many sectors in traditional agricultural markets. Locating a sorghum-based ethanol plant introduces a new demand source for sorghum as feedstock, therefore distorting the local demand and supply equilibrium. Intuitively, it is expected that the new plant increases the local demand for sorghum and, consequently, leads to a rise in local sorghum prices (change in basis). As a result, it is likely that higher prices of sorghum provide an incentive for local farmers to convert more acres planted to sorghum in the vicinity of the plant at the expense of other crops, mainly cotton, where a large local infrastructure exists. Thus, studying the ethanol plant impacts could shed light on the issue about the relationship between ethanol production and regional economic activity related to the cotton industry.

Traditional models of the effects of ethanol plants fail to consider the spatial distribution effects. That is, due to transportation costs, it is expected that the ethanol plant has a greater impact on spatially closer cropland compared with more distant farms. The spatial distribution is important to understand to anticipate how changes in demands on infrastructure (cotton gins, for example) will be altered with the change in spatial distribution of acreage choices. Ignoring spatial dependence in empirical studies not only affect the magnitudes of the estimates and their significance, but may also give rise to serious errors in the interpretation of regression results. Thus, spatial models are needed to evaluate the effects of ethanol production on the local and regional economies. Further, using a spatial model will also allow for the differentiation between short- and long-run impacts which is important if we are to understand whether the plant location permanently or only temporarily disrupts the production decisions of local producers relative to the *status quo*.

The objective of this study is to examine the sorghum basis and cotton acreage distribution changes in Hockley County and its neighboring counties resulting from introducing a sorghum-based ethanol plant to the High Plains of Texas. To achieve this objective, this study models the sorghum basis and cotton acreage planted from 2002 to 2014 using county-level panel data collected from Hockley County that currently has a 40-million-gallon-per-year sorghum-based ethanol plant in operation, and its neighboring counties. Due to the spatial pattern of sorghum basis and cropland around an ethanol plant, the presence of a positive spatial autocorrelation is expected. Alternative spatial econometric models—a spatial Durbin model, a spatial lag model, and a spatial error model—are employed to account for any spatial dependence.

Study Area

The cotton acreage study area is the nine-county region of the Texas High Plains shown in Figure 1 and consists of the counties of Bailey, Cochran, Hale, Hockley, Lamb, Lubbock, Lynn, Terry, and Yoakum. These counties represented 32% of the total cotton production in Texas and 11% of U.S. cotton production from 2002 to 2014 (U.S. Department of Agriculture (USDA), National Agricultural Statistics Service (NASS), 2016). Cotton is a major industry and contributor to the economy of Hockley County and its surrounding region (the Texas High Plains). And, it is the most important crop in the region in terms of both acreage and crop value. An annual average of approximately 265,000 acres of cotton were planted in Hockley County for the period from 2002 to 2014, accounting for about 87% of the county's total cropland acreage (NASS, 2016).

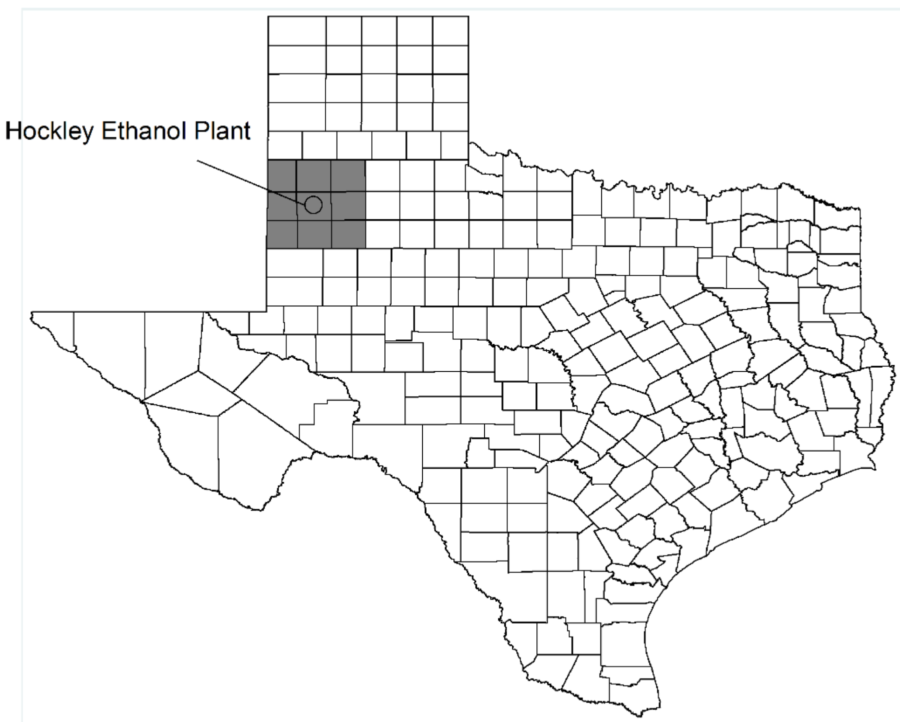


Figure 1. Cotton Acreage Study Area in Texas.

This study focuses on Hockley County and its surrounding areas because it is the location of the Levelland-Hockley County Ethanol plant, which began operations in 2008. The plant currently produces 40 million gallons of ethanol per year using only sorghum as a feedstock obtained from local producers. Assuming operating at full production capacity, the plant requires approximately 15 million bushels of sorghum per year (Guerrero et al., 2011). At the recent average state yield of 58 bushels per acre (NASS, 2016), this capacity is equivalent to approximately 260,000 acres of sorghum, or 12% of the 2.25 million acres of sorghum harvested in the state in 2014. The region is known for cotton production, while sorghum is considered primarily as a second crop planted behind failed cotton or in a planned rotation. Because ethanol production increases the demand for sorghum and is predicted to increase the local sorghum price, it allows local farmers to incorporate sorghum into their crop rotation with cotton to maximize their revenues.

Shifts in the agricultural sector are expected in neighboring communities due to increasing demand of sorghum for ethanol production. To the extent that ethanol production affects farmers' cropping decisions, *ceteris paribus*, it is important to quantify how the opening of an ethanol plant causes farmers to alter those choices to provide more information in expected changes to regional infrastructure demands.

Literature Review

Several studies have examined the effects of ethanol production on both local and national corn prices. Early work measuring the impact of ethanol plants on local basis patterns was conducted by McNew and Griffith (2005). With a data set of 12 ethanol plants that opened from 2001 to 2002, they found that the corn basis around a plant rose by 5.9 cents per bushel on average. Behnke and Fortenbery (2011) also examined the impacts of local ethanol plants on corn basis. Their results indicated that ethanol production within a 50-mile region of a county centriod has a small yet positive impact on local corn prices. And a 50 -million-gallon -per -year plant led to a 0.4 cent per bushel increase in corn basis. Contrary to previous findings, Katchova (2009) analyzed the spatial effect of ethanol biorefinery locations on local corn prices and concluded that ethanol plants have no significant effect in raising corn prices for farmers located close to ethanol plants, while prices in real terms have risen over time. However, few researchers have directly addressed the impacts of ethanol plants on cropland distribution changes. More relevant to the current study is the impact of ethanol production on local agricultural land use.

Turnquist, Fortenbery, and Foltz (2008) examined the effects of corn ethanol production on local land use and residential land values using data from 2000 to 2006 in Wisconsin. Their analysis considered whether agricultural land use trends are different in areas where agricultural production contributes to an ethanol plant's feedstock source compared to areas that are outside the purchase range of an ethanol plant. Their results showed that the agricultural land conversion was not affected by ethanol plants in their proximity. They explained that the positive commodity price effects resulting from the ethanol production are not so large as to influence farmers' choices. Secchi et al. (2011) examined the impact of the biofuels industry in Iowa on both current cropland and land in the Conservation Reserve Program (CRP) and its environmental consequences. In their analysis of land use change associated with the expansion of biofuels over the period from 2002 to 2006, the authors found that as corn prices increase, more cropland is planted with continuous corn because corn becomes relatively more profitable than soybeans. They concluded that substantial shifts in rotations favoring continuous corn rotations are likely if high corn prices are sustained. Finally, a recent paper closer to our study was conducted by Fatal and Thurman (2014). The authors examined the effect of U.S. ethanol production growth on surrounding corn acreage between 2002 and 2007. The results of this study implied that corn supply responds positively and locally to the changes in demand coming from the ethanol industry. Locating an ethanol plant in a certain area does trigger additional planted corn, especially locally around the plant. Moreover, the building of a typical plant was estimated to increase corn in the county by over 500 acres and to increase acreage in surrounding counties up to almost 300 miles away.

Most of these studies have focused on crops such as corn and soybeans, not cotton and sorghum. And the findings of these studies seem to indicate that the local grain basis and land use change near ethanol plants can be quite variable based on production size and location. In contrast with existing studies, we focus on the impact of locating an ethanol plant on the sorghum basis and cotton acreage distribution using a spatial structure, which is unique in the literature. Given the important role played by cotton production in the Texas High Plains and the limited literature available, a comprehensive analysis of ethanol plant impacts would be useful in providing accurate information for policy makers and other stakeholders.

Empirical Model and Data

The literature in spatial models is divided about whether to apply the specific-to-general approach or the general-to-specific approach (Florax, Folmer, and Rey 2003; Mur and

Angulo, 2006). More generally, Elhorst (2014a) proposed a strategy that mixes both approaches. First, the non-spatial model is estimated to test it against the spatial model (specific-to-general approach). In case the non-spatial model is rejected, the Spatial Durbin Model (SDM) is estimated to test whether it can be simplified to the spatial lag or the spatial error model (general-to-specific approach). We start with the estimation of the linear regression model using Ordinary Least Squares (OLS). Even though the OLS model is often rejected in circumstances where the spatial interaction is present, its results can serve as a benchmark.

The OLS Model

This study used the local sorghum basis to reflect local supply and demand conditions. The basis is defined in this analysis as the local cash price minus the nearby futures contract price. The local cash price typically refers to the spot cash price in a specific local market at a particular time. The futures contract price represents a nearby futures contract price for the given commodity, determined by the Chicago Board of Trade. However, since there is no sorghum futures market, the sorghum basis was calculated using the corn futures market (Wilson and Lutgen, 2004). In this context, sorghum basis is defined as follows:

(1) *Local Sorghum Basis = Local Cash Sorghum Price – Corn Futures Price*

Consequently, the basis is negative if the futures price is above the cash price; the basis is positive if the futures price falls below the cash price.

The observed sorghum basis around an ethanol plant is considered as the outcome of the interaction between sorghum supply and demand in this area and the interaction with other cash markets in neighboring areas. The local grain supply and demand, transportation costs, and storage costs are a few of the many factors that have been identified to affect the basis (Adjemian et al., 2011; Behnke and Fortenbery, 2011; McNew and Griffith, 2005; Olson, Klein, and Taylor, 2007). In order to estimate the impact of ethanol production on local sorghum basis, it is important to control for these factors. Specifically, it is modeled as a linear function,

(2) $$(Basis)_{i,t} = \alpha_0 + \alpha_1*(Sorghum\ Production)_{i,t-1} + \alpha_2*(Diesel\ Prices)_t + \alpha_3*(Animal\ Unit)_{i,t} + \alpha_4\ (Dummy\ Variable)_t + \varepsilon_{i,t}$$

where $i = 1 \ldots N$, $t = 1 \ldots T$. The variable $(Sorghum\ Production)_{i,t-1}$ is the sorghum production for county i in year t-1; $(Diesel\ Prices)_t$ is the annual Texas diesel retail price in year t; $(Animal\ Unit)_{i,t}$ is the quantity of grain-consuming animals on feed for county i in year t; $(Dummy\ Variable)_t$ indicates the period from 2008 to 2014, which is when the ethanol plant was in operation; and $\varepsilon_{i,t}$ is assumed to be a vector of independent and identically distributed (i.i.d.) error terms.

Note that sorghum production is one-year-lagged values. This is due to the fact that farmers making their planting decisions in year t are affected by sorghum production from the previous year (in year t-1). Assuming the acreage is held constant, the volume of production is a variable that proxies for both the previous profitability of sorghum as well as the impacts of previous weather events on overall production (risk). There is less information available about profitability on a county-level basis, and production also captures more county-level effects of weather and other systemic factors that can vary from year to year (that is, that would not necessarily be captured by the fixed effects portion of a panel model). Thus, we use past production as a control for several variables simultaneously. It is expected to have a negative impact on sorghum basis. As supply increases, the local sorghum cash price falls relative to the corn futures price. The average Texas diesel retail prices are used as a proxy to account for the transportation costs, which represent the amount used to ship produced grain to the terminal market. It is considered that feedgrain sales revenues increase even more for local farmers with a nearby ethanol plant because they receive higher regional prices due to lower transportation costs (McNew and Griffith, 2005). The quantity of grain-consuming animals on feed is used to reflect local sorghum consumption for its other primary use. Several studies have estimated the empirical significance of grain consumption in the basis determination (Adjemian et al., 2011; Olson, Klein, and Taylor, 2007). Increased demand for sorghum in a region is commonly expected to strengthen price relationships in local grain markets. The dummy variable for the period of 2008 to 2014 is specified to reflect the impact of the establishment of an ethanol plant on cotton acreage relative to the base period of 2002 to 2007.

To determine the impacts on cotton acres planted resulting from an ethanol plant, the following model is specified to represent the relationship between the cotton acreage planted in each of the nine counties from 2002 to 2014 and explanatory variables, which include cotton production, cotton futures price, the interest rate, and a dummy variable.

$$\begin{aligned}(3)\quad \log(Cotton\ Acreage)_{i,t} &= \alpha_0 + \alpha_1 {*} \log(Cotton\ Production)_{i,t-1} + \alpha_2 {*} \log(Cotton\ Future)_t \\ &\quad + \alpha_3 {*} \log(Interest\ Rate)_t + \alpha_4 {*} \log(Dummy\ Variable)_t + \varepsilon_{i,t},\end{aligned}$$

where $i = 1 \ldots N$, $t = 1 \ldots T$. The variable *(Cotton Production)* $_{i,t-1}$ is the cotton production for county i in year t-1; *(Cotton Future)*$_t$ is the average futures price in year t; *(Interest Rate)*$_t$ is the annual rate in year t; and $\varepsilon_{i,t}$ is assumed to be a vector of i.i.d. error terms.

The lagged cotton production is expected to be positively related to cotton acreage planted in each county. That is, when acreage is held constant, higher production last year indicates both higher profitability and better weather, therefore leading to an increased liklihood of planting this year. The interest rate is included to account for storage costs. Behnke and Fortenbery (2011) used the interest rate as a proxy for storage costs in their study because it mimics opportunity costs. Specifically, if the opportunity cost of storage is high, farmers are expected to lower their supply by reducing acres planted to cotton.

The sorghum basis and cotton acreage observations in a region are likely to be affected by other explanatory variables observed in neighboring regions, which is due to the spatial dependence commonly detected in such estimations. As such, use of conventional estimation methods that omit the presence of spatial effects, like OLS, may not only affect the magnitudes of the estimates, but also their significance (Anselin, 1988). Therefore, to account for any potential spatial dependence present among observations, alternative econometric models that incorporate spatial effects are considered: an SDM, a spatial lag model, and a spatial error model.

The Spatial Regression Models

The SDM (Anselin, 1988) contains spatial lags of the dependent variable as well as spatially lagged values of the independent variables.[2] This model allows for indirect spatial interactions. In other words, the exogenous explanatory variables influence not only the dependent variable within their own county, but within neighboring counties as well. While the SDM was originally formulated for cross-sectional data, its specification and estimation has been extended based on spatial panels (Elhorst, 2014). The SDM can be expressed as:

$$(4) \qquad y_{it} = \alpha + \beta x_{it} + \delta \sum_{j=1}^{N} w_{ij} y_{jt} + \theta \sum_{j=1}^{N} w_{ij} x_{jt} + \mu_i + \lambda_t + \varepsilon_{it}$$

for $i = (1 \ldots N)$ counties over time periods $t = (1 \ldots T)$. Furthermore, y_{it} represents the dependent variable at i and t; x_{it} is a $k \times 1$ vector containing explanatory variables; δ, θ

[2] While the addition of a spatially correlated term would seem an improvement to the SDM model, Elhorst (2010) notes that the inclusion of all possible spatial interaction terms results in a model that is not identified.

measures the intensity of spatial interactions between neighboring counties; $\sum_{j=1}^{N} w_{ij} y_{jt}$ represents the interaction effects of the dependent variable y_{it} with the dependent variable y_{jt} in neighboring counties; w_{ij} is an element of the spatial weight matrix W that defines the relationship between any two counties i and j; $\sum_{j=1}^{N} w_{ij} x_{jt}$ represents the interaction effects of the independent variable x_{it} with the independent variable x_{jt} in neighboring counties; μ_i and λ_t are spatial- and time-fixed effects; and ε_{it} is error terms which follow the normal assumptions. The SDM demonstrates that the characteristics of a county and its neighbors are simultaneously considered in the analysis. This allows us to explore whether the cotton area in county i is related to the features of its neighbors and, if so, to answer how they are associated.

The SDM has been widely advocated as a starting point in an empirical study because it nests two popular, simpler models: the spatial lag model and the spatial error model. Besides the nesting nature, the SDM has other merits. According to LeSage and Pace (2009), the cost of ignoring spatial dependence in the dependent variable or in the independent variables is relatively high, which raises the problem of biased and inconsistent coefficient estimates. In such circumstances, LeSage and Pace (2009) argue that the SDM is the only means of producing unbiased coefficient estimates, even if the true data generation process is a spatial lag, spatial error, or combined spatial lag/spatial error model. Additionally, the spatial effects in the spatial lag and spatial error models do not capture both local and global spillover effects (Elhorst, 2014) which may undermine the understanding of why the outcome of interest is spatially correlated.

By imposing certain restrictions on the parameters of the SDM, the spatial lag model and the spatial error model can be regarded as special cases of the SDM. The spatial autoregressive model (SAR), or the spatial lag model, assumes that dependencies exist directly among the levels of the dependent variable y. That is, the acreage planted to cotton at location i is more likely to be influenced by the cotton acreage planted in neighboring locations. In particular, if $\theta = 0$ is imposed, the SDM reduces to the SAR model. Then, in the case where $\theta + \delta\beta = 0$, the SDM simplifies to the spatial error model (SEM), which accounts for the spatial interaction of the error terms. One reason for this might be that there are some spatially clustered factors that influence the dependent variable y but are omitted from the specification.

The SDM incorporates both the spatially lagged dependent and independent variables, and the endogeneity in the model makes the interpretations of the coefficient estimates different from traditional regression models. Specifically, each exogenous variable contains both a direct effect, the impact of the variable on the outcome within a county, and an indirect effect (the impact of the variable on the outcome in neighboring counties). Feedback effects are also present due to the inclusion of the spatially lagged dependent variable (LeSage and Pace, 2009). To see this, the impact of a change in an explanatory variable (X_r) on the outcome y across n spatial units in the study region is given by:

$$\partial y/\partial X_r = (I_n - \rho W)^{-1}(I_n\beta_r + W\theta_r) \tag{5}$$

where $\partial y/\partial X_r$ is an $n\times n$ matrix; I_n is the identity matrix; and β_r and θ_r represent the parameter estimates associated with the independent variable within a county and in neighboring counties, respectively. Equation 5 contains important implications for the interpretations of SDM coefficient estimates. That is, the change in the *rth* independent variable of a county will not only lead to the change in the dependent variable in the same county (direct impacts), but also affect the dependent variables in other counties (indirect impacts). LeSage and Pace (2009) propose summarizing the information contained in this matrix by averaging the diagonal elements of the matrix to determine the average direct impact of a change in the *rth* independent variable and averaging either the summed column or the row elements of the matrix (excluding the diagonal elements themselves) to determine the average indirect impact of a change in *r*.

The spatial weight matrix $W = (w_{ij}: i, j = 1, \dots n)$ defines neighbors, as well as the spatial relationships that exist among n geographic units. Thus, it is employed to reflect the structure of potential spatial interaction. It is a positive matrix, and each spatial weight, w_{ij}, is defined to reflect the spatial influence of location j on location i. Typically, the definition of neighbors used in the weights matrix is based on a notion of distance decay or contiguity. By convention, the diagonal elements of the weights matrix are set to zero, $w_{ij} = 0 \ \forall\, i = j$, and row elements are standardized such that they sum to one. There are numerous ways to construct a weight matrix, but there is no direct method of choosing one over another (Anselin, 2002). In the case of sorghum basis, the spatial weights matrix W is weighted by inverse distance weight $1/d$, where d is the distance between any two cash markets. This approach assumes that geographically closer factors would be weighted stronger than more distant factors. Meanwhile for the cotton acreage estimation, W is specified as a binary contiguity matrix, where W is an $n \times n$ positive matrix that specifies the neighborhood set for each observation. It is defined in a way that contiguous units are assigned weights of 1, and noncontiguous units are assigned weights of 0. Contiguous units are known as neighbors. The weights are row-standardized so that all the elements of each row sum to one, that is, $w_{ij}^s = w_{ij} / (\sum_j w_{ij})$.

Data

The descriptive statistics for all variables in the analysis are summarized in Table 1. Data for county-level sorghum production, animal units, cotton acreage, and cotton production were obtained from NASS. The cotton futures prices were obtained from the BRIDGE database. The prime interest rate was obtained from the Federal Reserve Bank Statistical Release as a control for borrowing costs and the opportunity cost of storing cotton. The average Texas diesel retail prices were obtained from the U.S. Energy Information Administration.

Table 1. Descriptive Statistic for Model Variables.

Variables	Min	Max	Mean	Std. Dev.
Basis (cents)	-0.92	0.46	-0.24	0.27
Sorghum Production (bushels)	0	9983000	2101641	1687784
Diesel Prices ($)	1.32	3.49	2.59	0.72
Animal Unit (heads)	12000	505000	265666.7	154209.9
Cotton Acreage (acres)	55600	345200	209200	77557.88
Cotton Production (480 lb bales)	12200	513000	209293.6	129941.2
Cotton Future (cents)	42.86	131.57	67.84	23.37
Interest Rate	1.41	5.27	3.51	1.25

To estimate sorghum basis changes resulting from the opening of an ethanol plant, daily local sorghum price and basis observations from several different grain markets in the Southern High Plains of Texas were purchased from Cash Bid Data Service (2017) from 2002 to 2014. For this study, annual sorghum basis series were created from the daily local sorghum basis observations from 15 different grain cash markets. Figure 2 illustrates the locations of these grain markets around the Levelland-Hockley County Ethanol Plant.

Figure 2. Sorghum Cash Markets Around the Levelland-Hockley County Ethanol Plant.

Results

Spatial Autocorrelation

At first, a global Moran's *I* test was used to examine spatial autocorrelation. First introduced by Moran (1950), the *I* statistic is the most commonly used measure of spatial autocorrelation, which is calculated as:

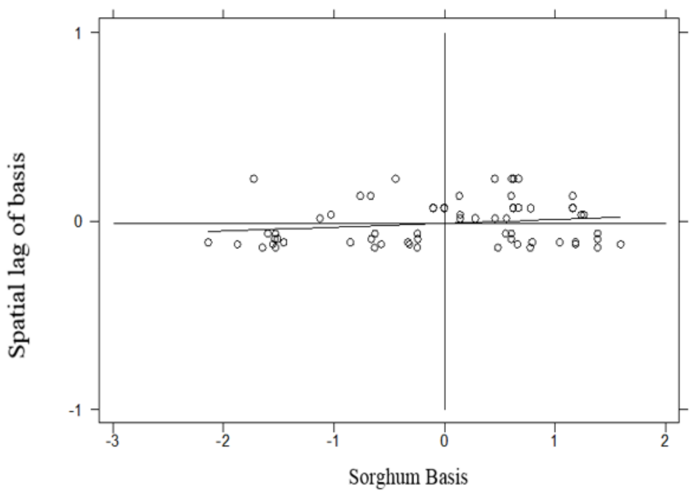
(6)
$$I = \frac{n}{S_0} \frac{\sum_i \sum_j w_{ij}(x_i - \bar{x})(x_j - \bar{x})}{\sum_i (x_i - \bar{x})^2}$$

where n is the number of locations; $\bar{x}$ is the mean of the x variable; w_{ij} are the elements of the weight matrix; and S_0 is the sum of the elements of the weight matrix $S_0 = \sum_i \sum_j w_{ij}$. The statistic varies from -1 to +1. A positive I value indicates that there is clustering of similar values across geographic space, while a negative I value indicates that neighboring values are more dissimilar.

Table 2 displays the results from the Moran's *I* test for both dependent variables to be estimated. The Moran's *I* statistics (0.02 and 0.05) are statistically significant; thus, the H_0 of no spatial dependence is rejected for both the sorghum basis and cotton acreage planted. To illustrate, Figure 3 and Figure 4 depict the Moran's scatter plot, which describes an observation's values in relation to its neighbors. The slope of the scatter plots corresponds to the value of Moran's *I*. As shown in Figure 3, counties with strong basis values are likely close to other counties with similar strong basis, while weak basis counties are likely to be surrounded by other weak basis neighbors. Likewise, counties with more acres planted to cotton are likely to be close to other counties with high cotton acreage, while counties with fewer cotton-planted acres are likely to be surrounded by other similar neighbors as shown by Figure 4. This sort of spatial dependency in cropping patterns is not unexpected given spatial patterns in climatic and soil type variables, among other factors.

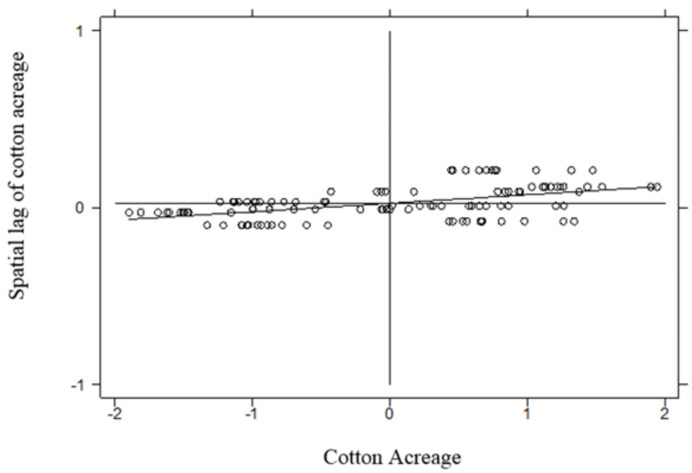
Table 2. Moran' *I* Test Results for Dependent Variables.

Variable	*I*	*E(I)*	*Sd(I)*	*Z*	*p-value*
Sorghum Basis	0.02	-0.02	0.02	2.17	0.02
Cotton Acreage	0.05	-0.01	0.01	9.75	0



Moran's *I* of Sorghum Basis: 0.02, *p*-value = 0.02

Figure 3. Moran's Scatterplot for Sorghum Basis.



Moran's *I* of Cotton Acreage: 0.05, *p*-value = 0.00

Figure 4. Moran's Scatterplot for County Level Cotton Acreage.

Model Results and Selection

Table 3 summarizes the estimation results for the sorghum basis. Four regression models are presented: the OLS model, SAR with fixed-effects, SEM with fixed-effects, and the dynamic SDM with fixed-effects, respectively. Spatial models were estimated using the maximum likelihoood method. Early empirical studies (Adjemian et al., 2011; Behnke and Fortenbery, 2011; McNew and Griffith, 2005) have shown that neighbors' characteristics may contribute to the basis values within a county. In addition, the significant spatial lag (ρ) and spatial error (λ) coefficients indicate the presence of spatial interaction. These results indicate that the conventional analytical approach fails to take the features of spatial pattern into account. Thus, OLS is not appropriate for the analysis in this situation. We adopted a likelihood ratio (LR) test as in LeSage and Pace (2009) and Elhorst (2014b) for model selection. The LR test is based on the log-likelihood function values of the different models. The test statistic follows an χ^2 distribution with degrees of freedom equal to the number of restrictions imposed. The LR test of the SDM versus the SAR model takes the value of 30.47 and is significant, which implies that the SAR model is rejected in favor of the SDMl. Likewise, the LR test of the SDM versus the SEM (30.82) indicates rejection of the SEM in favor of the SDM. Thus, the SDM is considered the preferred specification for the sorghum basis estimation. Hence, we concentrate on the interpretation of the SDM results. The spatial lag coefficient ρ in the SDM is significant and estimated to be 0.5, which implies that sorghum basis in area *i* covaries with basis values among its neighbors. Specifically, the coefficient indicates that sorghum basis in area *i* increases by about 0.5% when basis increases by 1% in surrounding areas.

Table 3. Models Estimation Results for Sorghum Basis.

	OLS	SAR	SEM	SDM
Intercept	-0.09			
Standard Error	*-0.09*			
Fixed-effects		0.01	-0.12	1.35
Lagged Sorghum Production	-1.74e-08*	-2.69E-09	-4.56E-09	-1.34E-08
	-1.01E-08	*-1.07E-08*	*-1.57E-08*	*-1.66E-08*
Diesel Prices	0.06	0.02	0.06	-0.42***
	-0.04	*-0.03*	*-0.08*	*-0.04*
Animal Unit	-2.25e-07**	-2.57E-07	-2.04E-07	1.13E-07
	-1.05E-07	*-4.02E-07*	*-3.74E-07*	*-2.41E-07*
Dummy Variable	-0.18***	-0.04	-0.17	0.67***
	-0.05	*-0.07*	-0.19	*-0.05*
ρ		0.65***		0.54***
		-0.08		-0.08
λ			0.65***	
			-0.08	

* Notes significance at 10%. ** Notes significance at 5%. *** Notes significance at 1%.

The coefficient estimates of the SDM model are not directly interpretable due to the feedback effects (LeSage and Pace, 2009) present among neighboring counties. By including lags of the dependent variables as well as lags of the independent variables in both space and time, the SDM provides the magnitude of direct and indirect effects, both in the short term and long term. The direct effects measure the change expected within any individual county, while the indirect effects measure the magnitude of spatial spillover effects accumulated over all counties in the study. The total effect is the sum of these direct and indirect effects. The estimation results of the direct, indirect (spatial spillover), and total effects are reported for each of the independent variables in Table 4. As indicated by the estimated effects, the significant indirect effects (spatial spillover effects) provided strong evidence to support our argument that the characteristics of neighboring counties are important determinants of sorghum basis. Three variables demonstrating significant spillover effects on sorghum basis are lagged sorghum production, diesel prices, and the dummy variable. Although with the expected sign, animal unit is found to be the only variable that shows insignificant results.

Table 4. Estimated Effects from Spatial Durbin Model for Sorghum Basis.

	Short-term			Long-term		
	Direct	Indirect	Total	Direct	Indirect	Total
Lagged Sorghum Production	-3.10e-08*	-3.92e-07***	-4.23e-07***	1.12E-08	9.33e-08**	1.04e-07***
	-1.83E-08	*-8.91E-08*	*-9.97E-08*	*-3.12E-08*	*-3.30E-08*	*-5.63E-09*
Diesel Prices	-0.44***	-0.49**	-0.93***	-0.75***	0.98***	0.23***
	-0.05	*-0.2*	*-0.24*	*-0.07*	*-0.08*	*-0.02*
Animal unit	1.17E-07	1.38E-07	2.55E-07	2.02E-07	-2.63E-07	-6.08E-08
	-2.46E-07	*-2.93E-07*	*-5.34E-07*	*-4.24E-07*	*-5.52E-07*	*-1.29E-07*
Dummy Variable	0.70***	0.78**	1.48***	1.21***	-1.57***	-0.37***
	-0.06	*-0.29*	-0.32	*-0.009*	*-0.11*	*-0.03*

* Notes significance at 10%. ** Notes significance at 5%. *** Notes significance at 1%.

It is expected that more sorghum produced locally is associated with a weaker local basis as the local price falls with increasing supplies. This is confirmed by our results. The short-run direct and indirect effects (spatial spillover effects) for lagged sorghum production is negative and statistically significant. This indicates that an increase in sorghum production will lead to a short-run decrease in sorghum basis not only in that county itself, but also in neighboring counties. Similar effects were found for the variable of diesel prices, which acts as a proxy for transportation costs in the model. More specifically, the short-run total impact of diesel prices is negative and statistically significant, indicating that a $1.00 increase in the price of diesel will cause the basis to drop by 0.9 cents per bushel overall. At last, it is worth mentioning that the short-term total effects are higher in absolute values than the long-term counterparts for both lagged sorghum production and diesel prices, 4.23e-07 and 0.93 in the short-run, as compared to 1.04e-07 and 0.23, respectively, in the long-run. This result simply indicates that the sorghum basis is more sensitive to short-run circumstance changes.

Finally, the the Levelland-Hockley County Ethanol Plant, as reflected in the estimates of the dummy variable, has both positive direct and indirect effects (both are statistically significant) on the sorghum basis in the short run. The coefficients suggest that, since the ethanol plant has been established, the sorghum basis rose by 0.7 cents per bushel in Hockley County. And, it also leads to a short-run increase in the sorghum basis by 0.8 cent in neighboring counties. The short-term total impact of the dummy variable is also positive. This total impact includes the direct effect of ethanol production on the sorghum basis within a county, as well as the indirect effect from the ethanol plant in all other neighboring counties. Therefore, the total effects provide relevant estimates from a broader perspective for policy makers and other partcipants in the market. The result suggests that locating an ethanol plant will cause the basis to rise by 1.5 cents per bushel overall. While in the long run, the positive direct effect of 1.2 from ethanol production are offset by the negative effect of -1.6 on the sorghum basis from spatial spillover effect, leading to a total effect of -0.4 in the long run. Thus, it appears that the ethanol plant actually decreased sorghum basis over the whole region in the long run, but the economic magnitude is small.

Table 5 shows the estimation results of cotton acreage based on four regression models (OLS, SAR, SEM, and SDM). When comparing different models, the main variable of interest, the dummy variable, exhibits the expected negative sign and are statistically significant in most cases except for SEM. On the other hand, the one-year lag in cotton production was found to be positively related to the cotton acres planted in the conventional approaches. However, the SDM results did not provide statistical support for this finding. More generally, this result indicates that, after controlling for spatial effects, the lagged own-county production level does not influence planted acres in the current period.

The spatial lag (ρ) and spatial error (λ) coefficients are statistically significant across three spatial models, indicating the presence of spatial interaction. Again, an LR test is applied to determine the appropriate specification. The LR test of the SDM versus the SAR model takes the value of 4.68 and is significant, which implies that the SAR model is rejected in favor of the SDM. Likewise, the LR test of the SDM versus the SEM (6.12) indicates rejection of the SEM in favor of the SDM. Thus, the SDM is considered the preferred specification for the cotton acreage model. Since the diagnostic test results suggest that the SDM provides a better fit, we will limit our interpretation to the SDM results. The spatial lag coefficient ρ in the SDM is significant and estimated to be 0.5, which implies that cotton acres planted in area *i* covaries with the cotton acreage among its neighbors. Specifically, the coefficient indicates that cotton acreage in area *i* increases by about 0.5% when cotton acres increase by 1% in surrounding areas.

Table 5. Models Estimation Results for County Level Cotton Acreage.

	OLS	SAR	SEM	SDM
Intercept	3.37***			
Standard Error	*-0.43*			
Fixed-effects		2.2	5.03	4.54
Lagged Cotton Production	0.47***	0.04***	0.05***	-0.04
	-0.04	*-0.01*	*-0.01*	*-0.02*
Cotton Future	-0.05	0.03	0.09	0.12**
	-0.19	*-0.03*	*-0.06*	*-0.06*
Interest Rate	-0.71***	-0.12**	-0.21*	-0.24**
	-0.17	*-0.05*	*-0.12*	*-0.12*
Dummy Variable	-0.15***	-0.03*	-0.06	-0.08*
	-0.04	*-0.02*	-0.05	*-0.04*
ρ		0.54***		0.51***
		-0.08		-0.07
λ			0.55***	
			-0.08	

* Notes significance at 10%. ** Notes significance at 5%. *** Notes significance at 1%.

Table 6 reports the direct, indirect, and total effects from the SDM for each of the independent variables. Overall, the short-term direct effects appear to be smaller than the long-term direct effects: -0.02 versus -0.23 for the lagged cotton production, 0.13 versus 0.43 for the cotton futures price, -0.27 versus -0.86 for the interest rate, and -0.09 versus -0.29 for the dummy variable. This is consistent with microeconomic theory because increases in time generally lead to more elastic responses. Because all variables in the model are log-transformed, the direct, indirect, and total effects can be explained in elasticity terms.

Table 6. Estimated Effects from Spatial Durbin Model for Cotton Acreage.

	Short-term			Long-term		
	Direct	Indirect	Total	Direct	Indirect	Total
Lagged Cotton Production	-0.02	0.21**	0.20**	-0.23*	0.33**	0.10**
	-0.02	*-0.1*	*-0.1*	*-0.12*	*-0.15*	*-0.05*
Cotton Future	0.13**	0.12*	0.25**	0.43**	-0.31*	0.13**
	-0.06	*-0.07*	*-0.12*	*-0.22*	*-0.17*	*-0.06*
Interest Rate	-0.27**	-0.25	-0.51*	-0.86**	0.60**	-0.26**
	-0.13	*-0.16*	*-0.28*	*-0.41*	*-0.3*	*-0.13*
Dummy Variable	-0.09*	-0.08	-0.17*	-0.29*	0.20*	-0.09*
	-0.05	*-0.06*	-0.1	*-0.15*	*-0.11*	*-0.05*

* Notes significance at 10%. ** Notes significance at 5%. *** Notes significance at 1%.

Both the direct and indirect effects are found to be positive and statistically significant for the cotton futures price in the short run. This indicates that an increase in the nearby futures price has a positive effect on cotton acres not only in that county itself, but also in neighboring counties. The total impact of the futures price on cotton acres is positive and statistically significant and suggests that a 1% increase in futures prices is associated with 0.25% increase in cotton acreage overall. Interstingly, the short-run elasticity is higher than the long-run elasticity. This result is not surprising and suggests that producers are responsive to short-run price changes, but, in the long run, both government programs and climatic conditions tend to limit price responsiveness.

Although the short-term direct impact of lagged cotton production on cotton acres is not statistically significant, the indirect impact is found to be positive and significant. The results in Table 6 indicate that the relationship between the lagged cotton production and cotton acres within a county (short-term direct) is only about 8% as strong as the spatial spillover effect (0.018/0.213=0.08). More importantly, the total effect for lagged production is positive and significant, suggesting that a 1% increase in lagged cotton production in a county will increase cotton acreage by 0.2% overall.

The interest rate acts as a proxy for the opportunity cost of storage as well as the cost of borrowed or own capital. When the opportunity cost of storage is high, it is likely for farmers to lower supply by decreasing cotton acres planted. As expected, it is found to be negatively related to cotton acreage for both the short-term direct and indirect effects. However, the indirect impact is not statistically significant. That is to say, a change in the interest rate will only affect the cotton acres within a county. On the other hand, a positive and significant spatial spillover effect is found in the long run.

Finally, for the variable of most interest, the dummy variable that indicates the presence of the ethanol plant, the analysis suggests a negative impact on cotton acreage in the short run for both direct and indirect effects. This finding confirms our hypothesis that installing an ethanol plant not only decreases cotton acres in vicinity of the plant, but also encourages farmers to convert more cotton acres planted to sorghum in neighboring counties. Furthermore, the spatial spillover effect is smaller than the direct effects (-0.08 vs -0.09), which makes sense because the impact of an ethanol plant is larger for the county where the facility is located. However, the spatial spillover effect of an ethanol plant is not statistically significant. Interestingly, the long-term direct effect of the dummy variable is negative with the spatial spillover effect being positive. Both effects are statistically significant. This result suggests that locating an ethanol plant in a county has a negative impact on cotton acreage in its own county, but has a positive impact on cotton acreages in neighboring counties in the long run. One explanation for this finding is that farmers located closer to the ethanol plant are more likely to be devoted to growing sorghum to satisfy the ethanol plant's demand for it as feedstock, mainly at the expense of cotton. Also, this result mirrors that of the basis (increase in the basis locally and decrease in the basis regionally in the long run).

The Levelland-Hockley County Ethanol Plant is a small plant, which only produces 40 million gallons of ethanol per year, or the equivalent of 15 million bushels of sorghum per year if operating at full production capacity. In the long run, as sorghum production from the county where the ethanol plant is located satisfies the facility's demand, the local basis for sorghum in the surrounding counties likely then returns to traditional levels that are more favorable to cotton production. Again, the impact of the ethanol plant on cotton acres is found to be greater locally (-0.29 vs 0.20). As a result, the total effect of the dummy variable appears negative and statistically significant, but the economic significance is small given that the total impact of the ethanol plant is a 0.1% reduction in total regional cotton acreage.

Conclusions

The purpose of this study was to demonstrate impacts on the sorghum basis and local cotton acreage distribution from a locally owned 40-million-gallon sorghum-based ethanol plant located in Hockley County, Texas, over the period from 2002 to 2014. The

Moran's *I* tests provided evidence of the existence of spatial dependence in the sorghum basis and cotton acreage, and LR tests were performed to determine the appropriate regression model. As suggested by the test statistics, the spatial Durbin model performs significantly better than the spatial lag and spatial error models. The SDM used here is a useful tool in research related to the spatial effects of the ethanol industry, as the explicit incorporation of these spatial interaction effects results in more informative estimates of the impacts of the explanatory variables. In particular, the short-run total impact on the sorghum basis resulting from hosting a 40-million-gallon ethanol plant was estimated to be a 1.5 cent rise per bushel overall. Moreover, the negative and significant total impact of the dummy variable indicates that a short-term average of 0.2% increase in cotton acres over all counties was assocaited with the introduction of the ethanol plant.

Our empirical results suggest that the sorghum basis and cropping patterns were affected by the introduction of an ethanol plant, facilitated in part by the favorable biofuel policies and the increased profitability of growing sorghum relative to cotton. Changes in the RFS affect ethanol production which, in turn, will affect local grain prices and farmers' revenues. Furthermore, the size and location of an ethanol plant plays a major role in affecting local sorghum basis and farmers' cropping decisions. In other words, rapid expansion in the ethanol industry can ripple through a regional economy from changing local supply/demand equilibrium all the way through to altering local cropland spatial distribution. Eventually, the effects of how those ripples expand and how far those ripples will extend vary widely depending on the size of the ethanol plant and the local economy structure.

Of greatest importance to both this particular case, as well as the broader understanding of the impacts of the location of ethanol plants, is that spatial effects do matter, especially in the long run. As we might expect, the introduction of a plant in the short run has immediate impacts as prices/decisions adjust to the new information and attempt to find a new spatial equilibirum. As time passes, the acreage adjustments become more concentrated in a local area (an increase in the elasticity), and the surrounding areas return to their previous patterns, and perhaps even compensate for the loss of acres in the affected county (ies). Thus, the overall impact of the plant is statistically important, but economically quite small in the long run.

Methodologically, however, these results suggest that analysis of the impacts of plant location like this requires attention to the spatial correlation of changes in the key variables. Simple comparison of the total long-run effect from the SDM shows that the OLS estimate was nearly twice as large as the SDM estimate, even assuming that one could trust the OLS estimates. Given that our effects were quite small due to the small size of the ethanol plant in our study, the impacts of larger plants is likely to exacerbate the problem of spatial correlation providing even stronger justification for the use of spatial methods.

To our knowledge, no study has attempted to examine the introduction of an ethanol plant on local cropland distribution using a spatial structure. By accounting for spatial dependencies, this study provided a greater understanding of the overall impact of ethanol production on local agricultural sector and confirmed the importance of the characteristics of neighbors with the spatial Durbin model. The method presented here provides an alternative approach that is complementary to existing spatial techniques and contributes to expanding the research related to the spatial effect of the ethanol industry on cropland changes. The model could be extended to include all ethanol plants within a state, which allows for a comprehensive study of the impacts of ethanol production on a state or regional level. Data used in the model could be continually updated to examine the long-term effects on local basis and cropland use changes.

References

Adjemian, M.K., T. Kuethe, V. Breneman, R. Williams, M. Manfredo, and D. Sanders. (2011, April 25-26). "A Spatial Approach to Estimating Factors that Influence the Corn Basis. " Paper presented at the NCCC-134 Conference on Applied Commodity Price Analysis, Forecasting, and Market Risk Management, St. Louis, Mo.

Anselin, Luc. (1988). *Spatial econometrics: Methods and models.* Dordrecht Boston: Kluwer Academic Publishers.

Anselin, Luc. (2002, November). "Under the hood: Issues in the specification and interpretation of spatial regression models." *Agricultural Economics* 27(3), 247–267.

Balat, M., and H. Balat. (2009, November). "Recent trends in global production and utilization of bio-ethanol fuel." *Applied energy* 86(11), 2273-2282.

Behnke, K., and T.R. Fortenbery. (2011, April 18-19). "The Impact of Ethanol Production on Local Corn Basis." *Proceedings of the NCCC-134 Conference on Applied Commodity Price Analysis, Forecasting, and Market Risk Management,* St. Louis, Mo.

Elhorst, Paul J. (2010, March). "Applied Spatial Econometrics: Raising the Bar." *Spatial Economic Analysis* 5(1), 9-28.

Elhorst, Paul J. (2014a). *Spatial Econometrics: From Cross-Sectional Data to Spatial Panels.* London: Springer.

Elhorst, Paul J. (2014b). Spatial panel models. *Handbook of regional science.* Springer, Berlin, Heidelberg.

Fatal, Y.S. and W.N. Thurman. (2014, May). "The Response of Corn Acreage to Ethanol Plant Siting." *Journal of Agricultural and Applied Economics* 46(2), 157–171.

Florax, R.J., H. Folmer, and S.J. Rey. (2003, September). "Specification Searches in Spatial Econometrics: The Relevance of Hendry's Methodology." *Regional Science and Urban Economics* 33(5), 557-579.

Gardner, Bruce. (2007, December). "Fuel Ethanol Subsidies and Farm Price Support." *Journal of Agricultural and Food Industrial Organization* 5(2), 1-22.

Guerrero, B., J. Johnson, S. Amosson, P. Johnson, E. Segarra, and J. Surles. (2011). "Ethanol Production in the Southern High Plains of Texas: Impacts on the Economy and Scarce Water Resources." *Journal of Regional Analysis and Policy* 41(1), 22-32.

Hill, J., E. Nelson, D. Tilman, S. Polasky, and D. Tiffany. (2006, July). "Environmental, Economic, and Energetic Costs and Benefits of Biodiesel and Ethanol Biofuels." *Proceedings of the National Academcy of Sciences* 103(30), 11206-11210.

Katchova, Ani L. (2009, July). "The Spatial Effect of Ethanol Biorefinery Locations on Local Corn Prices." Paper presented at the Agricultural and Applied Economics Association's 2009 Annual Meeting, Milwaukee, Wis.

LeSage, James, and R. Kelley Pace. (2009). *Introduction to Spatial Econometrics.* Boca Raton: Taylor & Francis Group.

Lewis, K.E., and G.T. Tonsor. (2011). "The Impact of Ethanol Production on Spatial Grain Market Relationships." *International Food and Agribusiness Management Review* 14(4), 201-213.

Linton, J.A., J.C. Miller, R.D. Little, D.R. Petrolia, and K.H. Coble. (2011). "Economic feasibility of producing sweet sorghum as an ethanol feedstock in the southeastern United States." *Biomass and Bioenergy* 35(7), 3050-3057.

McNew, K., and D. Griffith. (2005). "Measuring the Impact of Ethanol Plants on Local Grain Prices." *Review of Agricultural Economics* 27(2), 164-180.

Moran, P.A. (1950). "Notes on continuous stochastic phenomena." *Biometrika* 37(1/2), 17-23.

Mur, J., and A. Angulo. (2006). "The Spatial Durbin Model and the Common Factor Tests." *Spatial Economic Analysis* 1(2), 207-226.

Olson, A.L., N. Klein, and G. Taylor. (2007). "The Impact of Increased Ethanol Production on Corn Basis in South Dakota." Paper presented at the American Agricultural Economics Association Annual Meeting, Portland, Ore.

Renewable Fuels Association. (2017). "Building partnerships/growing markets: Ethanol industry outlook 2017." Renewable Fuels Association, Washington, D.C.

Rooney, W.L., J. Blumenthal, B. Bean, and J.E. Mullet. (2007). "Designing sorghum as a dedicated bioenergy feedstock." *Biofuels, Bioproducts and Biorefining* 1(2), 147-157.

Secchi, S., L. Kurkalova, P.W. Gassman, and C. Hart. (2011). "Land use change in a biofuels hotspot: the case of Iowa, USA." *Biomass and Bioenergy* 35(6), 2391-2400.

Turnquist, A., T.R. Fortenbery, and J. Foltz. (2008, July). "Progress or devastation? The effects of ethanol plant location on local land use." Paper presented at the American Agricultural Economics Association Annual Meeting, Orlando, Fla.

Wilson, R., and L. Lutgen. (2004). "Summary of Selected Grain Basis Patterns in Nebraska." Available at: http://digitalcommons.unl.edu/cgi/viewcontent.cgi?article=5776&context=extensionhist.