SPATIAL ANALYSIS OF U.S. ETHANOL PRODUCTION

Frank Tenkorang
University of Nebraska at Kearney
tenkorangf1@unk.edu

Deborah Bridges, Associate Professor
University of Nebraska at Kearney
Bridgesd@unk.edu

ABSTRACT

Site-specific factors are assumed to influence ethanol plant location and production. This paper aims at determining whether these factors actually have influence on plant size. The rapid expansion of the industry could make these factors crucial in its survival. Exploratory spatial analyses did not show spatial dependence among nearby ethanol plants. However, a spatial error model was found to be superior to its a-spatial version. The corn production is the only factor that has statistically significant influence on ethanol plant size. The implication of these results is beginning to be seen in the current transportation bottleneck facing the industry.

Keywords: ethanol, spatial analysis, corn, transportation, cattle

INTRODUCTION

The current surge in gasoline prices has prompted many countries to look for alternative sources of fuel. Ethanol production is touted by many as the harbinger of the United States' energy solution. Although ethanol can be obtained from many plant materials, current well developed technology favors corn-based ethanol. As expected, corn producing states are taking advantage of the potential of this infant ethanol industry to boost their economies. In addition to corn, many other factors have been identified to play significant role in the location of ethanol plants. These factors become crucial as the industry expands and competition among plants, for market share and inputs, intensifies. As of September 2007, there were 129 ethanol plants in operation, 76 under construction and 10 under expansion. Currently U.S. annual production capacity of ethanol is 6.9 billion gallons, which will approximately double when the plants under construction and expansion are completed. The rapid growth of the industry needs to be monitored closely to ensure that the factors influencing plant size and location are considered in order to meet the demands of this growing industry.

Most ethanol related studies range from energy requirements to discussions on feedstock requirements. For example, it is estimated that a gallon of ethanol will require $0.383 worth of natural gas, liquefied petroleum, and electricity (Hurt et al. 2006). According to the USDA, 56 percent of total U.S corn goes into feed/residual, 26 percent into food/seed/industrial, and the remainder is exported. Ethanol accounts for only 50 percent of the food/seed/industrial usage. As a result, the expansion of the industry will require corn to be transferred from other uses to ethanol if production does not increase significantly.

In addition to these requirements, site specific factors such as transportation and water quantity have been identified to have likely influence on ethanol plant location (CFDC, NEB, and USDA, 2006). An article by Krauss (2007) in The New York Times describes how the ethanol boom seems to be fading away because of distribution problems. While the influence of these site specific factors is perceived to be significant, there have been no empirical studies to this effect involving these factors and ethanol production to substantiate it. Such an empirical analysis will provide vital information for the industry as it expands. The goal of this study is to determine the influence these site specific factors have on ethanol plant size capacity, and for that matter on production level.

METHOD

An econometric model with ethanol production as the dependent variable and explanatory variables representing water availability, corn production, number of cattle in the area, distance to the nearest inter-state highway as independent variables. The model, as specified, is shown in Equation 1.

$$E_{j,i} = \beta_0 + \beta_1 C_i + \beta_2 H_i + \beta_3 R_i + \beta_4 K_i + e_i \quad (1)$$

where
$E_{j,i}$ = million gallons of ethanol per year produced by plant j in county i

Ci = thousand bushels of corn produced in the county i
Hi = available water in county i (irrigated area used as a proxy)
Ri = Distance of the nearest interstate highway to plant j
Ki = cattle herd (thousands) in county i
ei = residuals

Generally, among the three modes of transporting ethanol (barge, rail, truck), barge is most efficient followed by rail. However, for short hauls (less than 500 miles) truck seems to the most economic option (Thompson, 2006). This study assumed that more short hauls will be done since long hauls may make ethanol noncompetitive because of increased cost. As a result, distance to the nearest interstate highway was used instead of railway line.

In traditional econometrics using time series data, autocorrelation is present if successive error terms are dependent on preceding error terms. Uncorrected autocorrelation leads to inefficient coefficients and inflated significance level. If the econometrics involves spatial units then the traditional autocorrelation solutions using unidirectional lags are not appropriate. This is because spatial dependence, by nature, are two-dimensional and multidirectional (Florax et al., 2002).

As the ethanol industry expands, nearby ethanol plants are expected to have influence on production level of each other due to competition for resources and market shares. This type of influence is referred to as spatial effect which if present and ignored in a regression framework results in estimation bias and model misspecification (Anselin, 1988). Several spatial analysis tools are available to determine the presence of spatial effect.

Exploratory Analysis

There are two types of spatial effects: spatial dependence and spatial heterogeneity. Anselin (1988) defines spatial dependence as the existence of a functional relationship between what happens at one location (space unit) and what happens elsewhere. Spatial heterogeneity is due to the lack of uniformity among features of spatial units.

Exploratory spatial data analysis can be used to determine the presence of spatial effect. Figure 1 is a Choropleth map of 122 ethanol plants across the country. Production capacity of these plants ranges from less than 40 million gallons per year (mgy) to over 100 mgy. Most of the plants in the west have capacities of less than 50 mgy; hence, spatial lag could exist among those plants. The rest of the map is a mix of plants of all sizes; hence, spatial heterogeneity could be present.

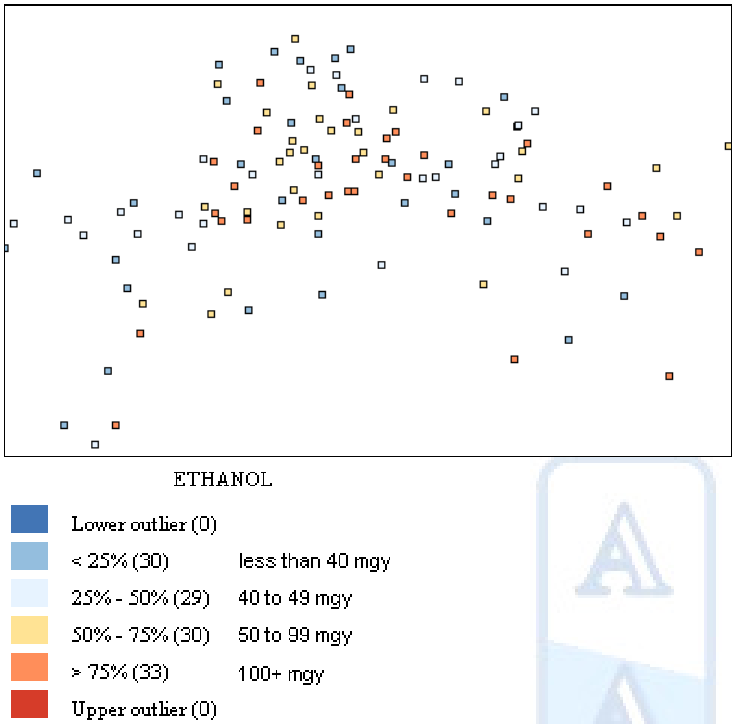

Figure 1 Choropleth Map of Ethanol Plant Size

The next exploratory tool is Moran scatter plot (Figures 2a and 2b). A Moran scatter plot shows observations in quadrants (Anselin, 1996). The first quadrant shows spatial dependence among large value observations while the third quadrant shows spatial dependence among small value observations. Having all or majority of observations in either quadrant is an indication of the presence of spatial dependence. A spatial heterogeneity is suspected if the observations are mainly distributed between quadrants I and II.

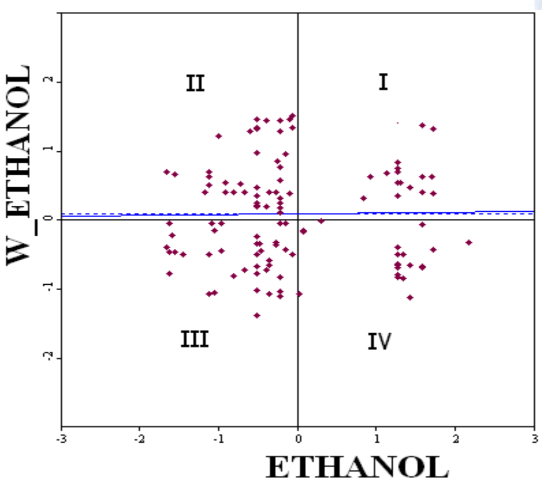

Figure 2a

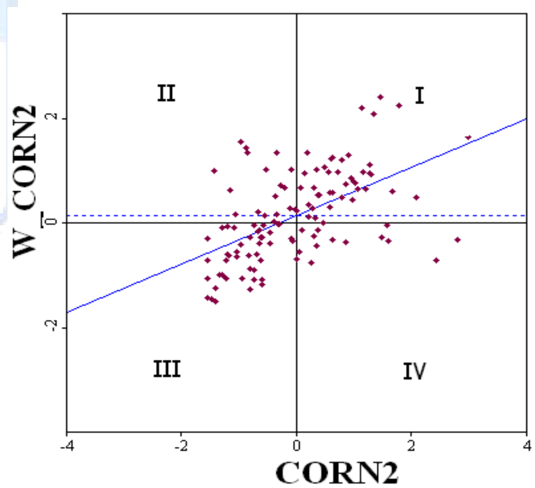

Figure 2b

The Moran scatter plot for ethanol shows that all the observations are within two standard deviations from the mean. Most of the observations are in the second and third quadrants thus indicating two main clusters. Quadrant II indicates a cluster of large and small plants, an indication of lack of spatial effect. Quadrant III indicates a cluster of small plant size, an indication that neighboring small plants have influence on one another. The corn Moran scatter plot shows two clusters in the Quadrants I and III. This is an indication of spatial dependence among nearby low corn producing counties, and also among nearby high corn producing counties. The two main clusters could also indicate spatial heterogeneity.

Exploratory analysis tools, such as those used above, are useful for diagnosing the presence of spatial effects. However, to determine which type of spatial effect (spatial dependence or spatial heterogeneity) is present, will require additional analysis through multiple regression.

Spatial Model

To determine the degree and type of spatial effect, tools such as Moran's I, Lagrange Multiplier spatial error and/or lag are used (Anselin, 1988). The two spatial effects are illustrated in the regression model below:

$$y = X\beta + \rho Wy + \lambda We + \varepsilon \qquad (2)$$

ρWy is the spatial dependence component

λWe is the spatial heterogeneity component

where y is the (n × 1) vector of the plant size; X is the (n × k) matrix of explanatory variables (Ci, Hi, Ri, and Ki); e is (n × 1) vector of OLS residuals from equation (1), β is (k × 1) vector of parameters; W is (n × n) spatial weight matrix capturing the influence of proximate ethanol plants' production. A distance weight matrix is used instead of a contiguous matrix due to the lack of contiguity among the plants. The two closest neighbors were selected for this purpose; ρ is spatial dependence parameter capturing the total influence of the dependent variables of neighboring spatial units, and λ is spatial heterogeneity parameter.

Data

Data was obtained from the following sources: County level shape files were extracted from ArcView GIS software version 3.3; Ethanol plant capacity data from Renewable Fuels Association; Corn production and cattle numbers from National Agricultural Statistics Service of United States Department of Agriculture (NASS, USDA); Area of irrigated land, obtained from NASS, USDA, was used a proxy for water availability; and Distance to the nearest inter-state highway was calculated using Google Earth 4.2.

RESULTS

The results of Moran's I test are presented in Table 1. Normality assumption is rejected for all variables hence randomization assumption is used in its stead. All the variables are spatially autocorrelated except ethanol.

Table 1 Results of Moran's I Test

		Randomization Assumption	
Variable	Wald Test*	Moran's I	Z-Value
Ethanol	8.07 (.017)	0.011 (0.72)	0.351
Corn	6.18 (0.045)	0.466** (<0.000)	8.886
Cattle	1320.48 (<0.000)	0.203** (<0.000)	4.23
Water	223.08 (<0.000)	0.572** (<0.000)	11.10

*Wald test rejects normality assumption for all variables at 5 percent level of significance
** Spatial autocorrelation is present at 1 percent level of significance

The statistically insignificant Moran's I for ethanol (dependent variable) suggests the absence of a spatial lag while the significant value for the independent variables suggests dependence among proximate counties with respect to corn, cattle and water. This could also indicate a spatial error model. Because of the inclusive spatial test, three models (a-spatial OLS, spatial lag and spatial error) were estimated for comparison. The results are presented in Table 2.

Table 2 Estimation Results of the Three Models

	OLS	*Spatial Error*	*Spatial Lag*
Variable	Coefficient	Coefficient	Coefficient
Constant	50.19* (< 0.000)	50.35* (<0.00)	52.66* (<0.000)
Corn	0.0006* (0.0221)	0.0006* (0.011)	0.0062* (0.0135)
Cattle	-0.0227 (0.687)	-0.034 (0.54)	-0.027 (0.627)
Water availability	-0.0142 (0.496)	-0.012 (0.54)	-0.014 (0.49)
Inter-State distance	-0.132 (0.749)	-0.156 (0.69)	-0.142 (0.72)
Lambda/Rho	---	-0.069* (<0.00)	-0.04 (0.615)
R^2 AIC	0.059 1204.0	0.071 1203.1	0.063 1205.6

*statistically significant at 5 percent level P values on parenthesis

Test for multicollinearity yielded a condition number of 4.95 (based on a-spatial model). Values less than 30 indicate multicollinearity is not a problem. The explanatory power of the model is substantially low (R^2 = 6.2%). Of the four site-specific factors known to influence the location of ethanol plants, only corn has significant influence on ethanol production (Table 2).

DISCUSSION

The results highlight vulnerabilities which are beginning to show up. Proximity of an inter-state highway is needed to ensure that ethanol reaches high fuel consuming areas on time. The current transportation bottlenecks in the industry seem to indicate that whatever the mode of transportation, it might have been passively considered in the plant size decision making. Most plants construction and expansion have been put on hold because of the transportation problems. The insignificance of cattle in the model also indicates an impending problem. The bulkiness of ethanol's main by-product, distiller's grain, requires that plants are located in areas with enough cattle to feed on it. However, the lack of significance seems to suggest plants plant developers do not consider the use of distiller's grain beforehand. Ethanol production requires about four gallons of water to per gallon. The results indicate that water is not a limiting factor because it is either in abundance or recycled. Corn, however, is the major factor in plant size consideration.

CONCLUSION

Spatial analysis of current ethanol plants reveals no spatial dependence among the plants. On the other hand, the spatial error model was superior compared to the a-spatial model. This means independent analysis of ethanol plants could lead to the inefficient results. Among the site specific factors considered, only corn production appears important in plant size consideration. The US consumes about 130 billion gallons of gasoline a year. The current ethanol production level is only about five percent of gasoline consumption. The potential of ethanol relieving the country of its dependence on gasoline is contingent on improvement in technology which will significantly increase ethanol to corn ratio, and also enable ethanol to be produced from other materials. The insignificance of the examined site-specific factors will reverse as production increases and competition among plants increases. Therefore, to prevent future problems such as the current clog in transportation, it is imperative that plant sizes are chosen in consideration to relevant site-specific factors.

REFERENCE

Anselin, L. (1988). *Spatial Econometrics: Methods and Models*. Dordrecht: Kluwer Academic Publishers.

(1996). The Moran Scatterplot as an ESDA Tool to Assess Local Instability in Spatial Association. In M. Fischer, H. Scholten, and D. Unwin (eds.), *Spatial Analytical Perspectives on GIS (*pp. 111–125) London: Taylor and Francis.

CFDC, NEB, and USDA (2006). A Guide for Evaluating the Requirements of Ethanol Plants. Retrieved October 27, 2007, from http://www.ethanol.org/pdf/contentmgmt/

Florax, J.G.M., R.L. Voortman, and J. Brouwer (2002). Spatial Dimensions of Precision Agriculture: A Spatial Econometric Analysis of Millet Yield on Sahelian Coversands. *Agricultural Economics*, 27, 425-443.

Hurt, C., W. Tyner, and O. Doering (2006). Economics of Ethanol. *Bio Energy*. Purdue University Cooperative Extension Service, ID-339. Retrieved December 6, 2007, from http://www.ces.purdue.edu/bioenergy

Krauss, C. (2007, September 30). Ethanol Boom Stalling as Glut Depresses Price. *New York Times*. Retrieved October 14, 2007, from http://www.nytimes.com

Thompson, S. (2006). Keep on truckin': Ethanol Boom Creates Transportation Challenges. *Rural Cooperatives*, 75(5). United States Department of Agriculture Rural Development. Retrieved on October 27 2007, from http://www.rurdev.usda.gov/rbs/pub/sep06/content.htm