Journal of Agricultural and Applied Economics, 46,2(May 2014):157–171

The Response of Corn Acreage to Ethanol Plant Siting

Yehushua Shay Fatal and Walter N. Thurman

U.S. ethanol production capacity increased more than threefold between 2002 and 2008. We study the effect of this growth on corn acreage. Connecting annual changes in county-level corn acreage to changes in ethanol plant capacities, we find a positive effect on planted corn. The building of a typical plant is estimated to increase corn in the county by over 500 acres and to increase acreage in surrounding counties up to almost 300 miles away. All ethanol plants are estimated to increase corn production by less than their annual requirements.

Key Words: bioenergy, biofuels, bootstrap, corn, energy, ethanol, renewable, spatial

JEL Classifications: Q13, Q15, Q16, Q18

Between February 2002 and February 2008, the number of corn-based ethanol plants in the United States grew from 49 to 129. Over the same period, annual ethanol production capacity increased from approximately two and a half billion gallons to almost eight billion gallons. These growth rates of 160% in the number of plants and 220% in ethanol production capacity were mainly the result of three factors: new energy regulation, a ban on the fuel additive methyl tertiary butyl ether (MTBE), and ethanol production incentives.

Changes in U.S. energy policy came primarily from the first and second Renewable Fuel Standards (RFS) acts. The RFS mandates have required gradually increasing ethanol use in the United States. The first RFS requirement was a part of the Energy Policy Act of 2005 and the second RFS was part of the Energy Independence and Security Act (EISA) of 2007. The stated goals of the EISA are to reduce U.S. energy dependence and greenhouse gas emissions. Whereas the first RFS served as the building block for an ethanol volume mandate, the second RFS expanded the first. Under the second RFS mandate, the fuel sector is subject to mandatory ethanol use of approximately ten billion gallons in 2008, increasing to over 35 billion gallons in 2022.

A different regulatory driver of ethanol use was a series of MTBE bans. MTBE, used as a fuel additive to reduce carbon monoxide emissions, was found to contaminate ground water and was banned in 19 states by 2004. As a result, a substitute chemical was needed to oxygenate gasoline and gasoline refiners began using ethanol as a substitute for MTBE.

Finally, federal and state incentives in the form of a tax credit and an excise tax credit for ethanol producers and blenders have played significant roles in the growth of the ethanol industry. These incentives increase the attractiveness of ethanol production and induce prospective ethanol producers to invest in production facilities. The Volumetric Ethanol Excise Tax Credit is the major federal incentive. It

Yehushua Shay Fatal is a senior research analyst at the University of North Carolina's Kenan-Flagler Business School, Chapel Hill, North Carolina. Walter N. Thurman is a Neal Reynolds Distinguished Professor in the Department of Agricultural and Resource Economics at North Carolina State University, Raleigh, North Carolina.

is a tax credit of 51 cents per gallon given to ethanol blenders against their fuel tax liability.[1]

Coincident with the increase in ethanol use and production has been an increase in the use of its principal feedstock: corn. The relationship between corn and ethanol is dual. In the first instance, the availability of corn attracts ethanol plants. As previous studies (Fatal, 2011a; Lambert et al., 2008) show, proximity to input markets is the most important determinant of ethanol plant location. Plants are attracted to corn-growing areas because corn accounts for 50–70% of ethanol input cost, depending on corn price as indicated by both The Guide for Evaluating the Requirements of Ethanol Plants and also Shapouri et al., (1998).

The shorter the distance between the corn source and the plant, the lower are transport costs and the smaller is the uncertainty of having corn in time for production. The attraction of ethanol plants to intensive corn production areas such as the Corn Belt[2] is demonstrated in Figures 1 and 2, showing the location of ethanol plants and the spatial distribution of planted corn.

A reverse causal chain also connects corn with ethanol. The siting of a plant induces a local increase in demand, which puts pressure on prices and induces an increase in acres planted to corn in vicinity of the plant. This second causal chain is the focus of the present article.

Understanding how ethanol plant siting affects the quantity and location of planted corn is important for several reasons. Knowing the local supply adjustment pattern of corn to the introduction of new ethanol plants can provide useful information to ethanol producers as they plan capacity and coordinate logistics. Furthermore, increases in future corn supply around the plant will support future plant expansion once market conditions encourage it. Other parties who may benefit from the article's results include ethanol industry investors and lenders, farmers who produce corn or buy Dried Distillers' Grain, and ethanol buyers. More generally, understanding the dual relationship between ethanol and agriculture is vital to an understanding of the increasingly important interface between the agricultural and energy sectors.

Industry Background

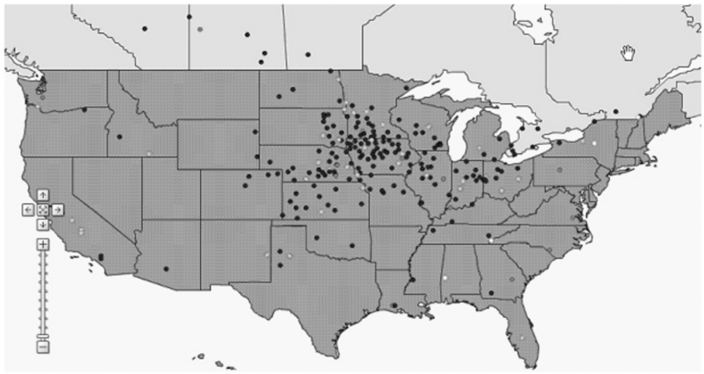
Corn's more traditional use is as a basic ingredient in the food industry and as an input into feed for cattle, hogs, and poultry. The United States is a net corn exporter, shipping corn overseas for livestock feed and other purposes. Before the rapid growth of the ethanol industry in the United States, most corn was sold for food production, feed, or export. After the significant increase in U.S. ethanol production, around 2002, production of corn increased. Not only did corn production increase, but the share of corn directed to the ethanol industry also rose significantly. Table 1 provides a detailed breakdown of U.S. corn production and use. It shows that corn production increased from approximately nine billion bushels in marketing year 2002–2003 to 13 billion in 2007–2008. The share of ethanol use in total production of corn rose from approximately 11% in marketing year 2002–2003 to over 23% in 2007–2008 and over 40% in 2010–2011.

Against the backdrop of this remarkable growth, it is notable that corn consumption by the ethanol industry is expected to continue to rise as mandated by the RFS. The second RFS sets a total U.S. ethanol production requirement of 36 billion gallons by the year 2022, where 15 billion gallons out of the total will be produced from corn by the year 2015.

The natural consequence of increased demand for corn is higher corn prices, which provide an incentive for farmers to allocate more land and other resources to growing corn at the expense of growing crops such as soybeans and wheat as well as drawing land from sources other than that previously allocated to soybeans and wheat. Consequently, the quantity produced of soybeans decreases and its price increases along with corn price. See Figure 3 for more details on growing more corn

[1] The incentive of 51 cents was established in the 2004 Job Bill. In 2008 Farm Bill, Congress reduced the incentive to 45 cents. The incentive expired in December 2011.

[2] The Corn Belt region mainly includes the states of Iowa, Indiana, Illinois, and Ohio. Also included are parts of South Dakota, North Dakota, Nebraska, Kansas, Minnesota, Wisconsin, Michigan, Missouri, and Kentucky.



Key:
- Producing
- Idle
- Under Construction

Source: Ethanol Producer Magazine, August 2009.

Figure 1. Ethanol Plant Location in 2009

and less soybeans. In addition to shifting land from other crops, additional lands not currently used for crop production can be brought into production. According to the U.S. Department of Agriculture (USDA) Economic Research Service–Corn Background, cropland used as pasture, reduced fallow, and acreage returning to production from expiring Conservation Reserve Program contracts are all land uses that were converted to corn production as a result of increases in corn price. However, it is likely that changes in land designation take more time to react to corn price increases than do changes in the crop rotation behavior of farmers.

One of the major agronomic considerations for a farmer is crop rotation. Without any special

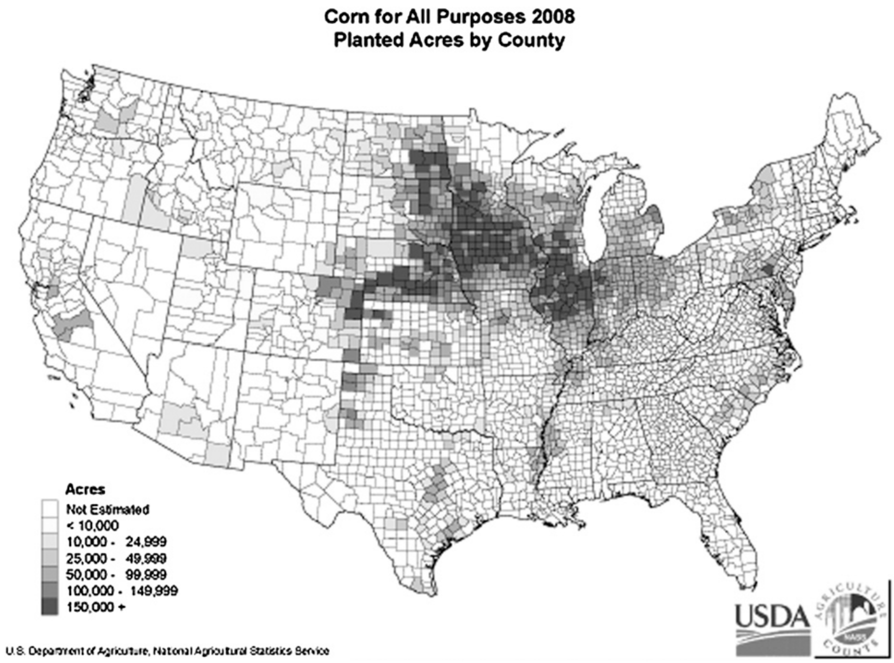


Figure 2. Planted Corn Acres in 2008

Table 1. U.S. Corn Production and Use, 2002–2008 (millions of acres and bushels)

	2002–2003	2003–2004	2004–2005	2005–2006	2006–2007	2007–2008	2008–2009	2009–2010	2010–2011
Acreage and yield									
Harvested area (million acres)	69.3	70.9	73.6	75.1	70.6	86.58	78.6	79.5	81.4
Yield (bushels/harvested acre)	129.3	142.2	160.3	147.9	149.1	150.7	153.9	164.7	152.8
Supply sources									
Production (million bushels)	8,967	10,087	11,806	11,112	10,531	13,038	12,092	13,092	12,447
Imports	14	14	11	9	12	20	14	8	30
Drawdown of stock	510	129	−1,157	147	663	−320	−50	−34	828
Total supply	9,491	10,230	10,660	11,268	11,206	12,738	12,056	13,066	13,305
Uses									
Feed and residual	5,563	5,793	6,155	6,152	5,591	5,938	5,182	5,140	5,000
Food, seed, and industry (excluding ethanol for fuel)	1,344	1,369	1,364	1,379	1,371	1,337	1,316	1,371	1,380
Ethanol for fuel	996	1,168	1,323	1,603	2,119	3,026	3,709	4,568	5,050
Total domestic use	7,903	8,330	8,842	9,134	9,081	10,301	10,207	11,079	11,430
Exports	1,588	1,900	1,818	2,134	2,125	2,437	1,849	1,987	1,875
Total use	9,491	10,230	10,660	11,268	11,206	12,738	12,056	13,066	13,305

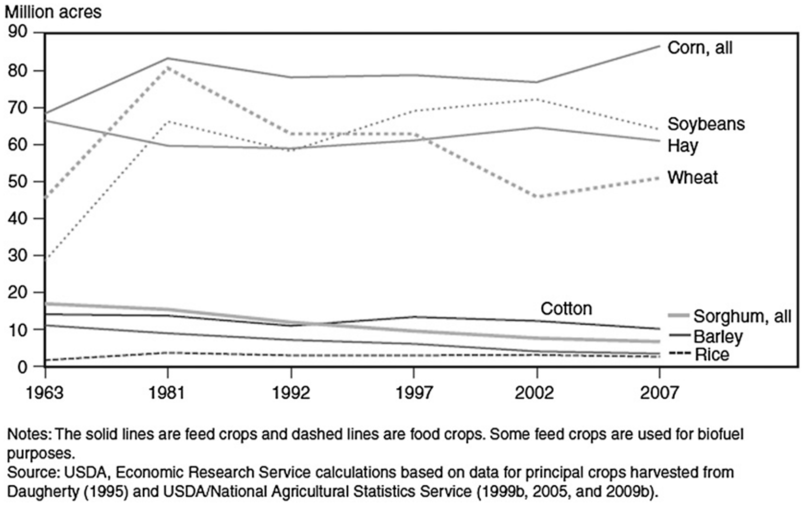


Figure 3. Principal U.S. Crops Harvested, 48 Contiguous States, 1963–2007

incentives to grow corn, farmers who grow soybeans and corn typically plant a year of corn followed by a year of soybeans. Growing corn one year after growing soybeans the previous year provides several advantages to farmers. First, corn yield is higher when farmers use a corn–soybean rotation. Second, using a corn–soybean rotation requires less tillage and less nitrogen fertilizer and pest control chemicals. However, instead of a corn–soybean rotation, farmers may opt to use a corn–corn–soybean rotation, two years of planting corn followed by one year of planting soybeans. Farmers choose a corn–corn–soybean rotation only if the market offers a sufficient incentive such as a higher corn price to offset the additional cost and lower yield that result from deviating from the corn–soybean rotation. Other important factors affecting a farmer's planting decision include expected future crop price and input costs, crop water needs, and weather forecasts that might affect yields.

A Spatial Econometric Model of Acreage Response to Ethanol Production Capacity

To answer the question of how ethanol plant siting affects surrounding corn acreage, we constructed a causal model that relates two key variables. The first represents the change in the capacity of the ethanol industry relevant to a particular county; the second represents the change in corn acreage caused by the change in the ethanol industry. Notice that the first variable requires the estimation of an ethanol plant's radius of influence on corn, which we embed in our structural model. Available data for corn acreage and ethanol capacity come at different levels of spatial resolution. Ethanol plants are discrete points in space (see Figure 1), whereas corn production and acreage are available at the county level. We represent the effective size of an ethanol plant by its rated capacity, which changes over time as plants are expanded.

There are several candidate variables one should use to represent corn (see Table 1). We choose planted corn instead of a production measure because it better represents farmers' intentions. Planted corn is a measure of input choice influenced by farmers' perception of ethanol-based demand for corn and less influenced by weather than is production. As an input choice, planted corn is also a function of other input prices, which we accounted for with time period dummy variables. With respect to corn supply response, we note that yields respond to price increases, which suggests that our calculated measures are likely conservative estimates of the corn supply response to ethanol plant siting.

To build an empirical model to identify the effect of the ethanol plant on corn acreage, we need to explicitly consider the nature of the data. We discuss that issue first before taking into account the time and space varying covariate that could, if not accounted for, confound the causal relationship we seek to identify in a county-level annual panel.

Given measures for ethanol capacity and corn acreage, consider the following spatial relation:

$$
\begin{aligned}
A_i &= \alpha + \beta Z_i + \varepsilon_i \\
&= \alpha + \beta \sum_{j=1}^{N} W(i,j) C_j + \varepsilon_i
\end{aligned} \tag{1}
$$

where:
A_i = planted corn for county i, measured in acres;
C_{jt} = ethanol production capacity of plant j, measured in millions of gallons per year;
$W(i,j)$ = a measure of the impact of plant j on county i's acreage; and
ε_i = an error term representing other determinants of acreage.

The weighted sum of ethanol plant capacity, labeled Z_i, in equation (1), is the effective capacity relevant to county i.

Equation (1) represents the "pull" from ethanol plants by weighing and summing the capacity for all plants, leaving unspecified the question of appropriate weights. One natural, albeit simplistic, weighting scheme would be to predetermine—or estimate—a radius distance from a county's centroid. Plants within the given distance from the county are "in." All others are "out." This leads to the following definition:

$$
\begin{aligned}
W(i,j) &= 1 \quad \text{if} \quad D_{ij} < r, \\
&= 0 \quad \text{otherwise},
\end{aligned} \tag{2}
$$

where D_{ij} is the distance from county centroid i to ethanol plant j.[3]

Equations (1) and (2) comprise a straightforward model, but they do not reflect the fact that the impact of an ethanol capacity should decrease with distance from the county centroid, and not discontinuously. Further away from the county, one would expect the effect of ethanol production to weaken; a 100-million gallon ethanol plant ten miles away from the corn source should have a greater impact on corn growers than the same plant 200 miles away.

A modification to the model can account for this effect by incorporating a linear relationship between plant–county distance and effective capacity.

$$
A_i = \alpha + \beta \sum_{j=1}^{N} max\left(1 - \gamma D_{ij}, 0\right) C_j + \varepsilon_i \tag{3}
$$

The term $max\left(1 - \gamma D_{ij}, 0\right) C_j$ is referred to as effective capacity. The capacity weight equals one when D_{ij} is equal to zero and diminishes to zero as D_{ij} increases. The value of γ is the rate at which effective capacity declines with distance. Figure 4 shows how effective ethanol plant capacity declines with distance. At a distance of $1/\gamma$, effective capacity declines to zero; plants farther from a county than $1/\gamma$ have no effect on corn acreage in the county.

An a priori assessment of $1/\gamma$, or the reach of an ethanol plant in its effect on corn acreage, can be obtained from previous research on the ethanol industry. AUS Consultants (2002) used a 50-mile radius in their analysis of economic benefits to a local community of building and operating an ethanol plant. Their reasoning for a using 50-mile radius was based on dry mill ethanol supply characteristics. These characteristics indicate that the vast majority of corn comes to plants from within a 50-mile radius to minimize grain transportation costs (U.S. Department of Agriculture, Economic Research Service, 2007).

However, the extent of a plant's economic reach should exceed the radius of its source region, because prices and plantings beyond the

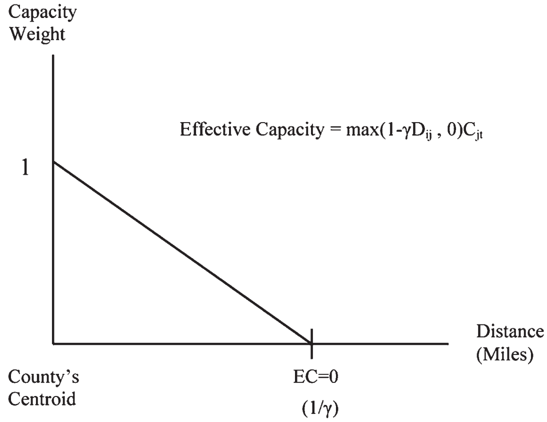


Figure 4. A Plant's Effective Capacity Declines Linearly with Distance from a County

[3] Here and in the empirical implementation of the model, we model D_{ij} as a straight-line distance. We view this as strongly correlated with, but ultimately an approximation to, the costs of transport between county i and ethanol plant j.

source radius will be influenced by changes in the supply–demand balance in the source region. This fact makes the dependence of effective capacity on distance a fundamentally empirical question. Accordingly, we estimate γ as a parameter of the model and derive effective capacity as

$$Z_i(\gamma) = \sum_{j=1}^{N} max\left(1 - \gamma D_{ij}\right) C_j$$

Now consider embedding the causal effect of effective capacity on county acreage in a panel model of county over time. Factors that should affect corn acreage in a given county include expected corn prices and input prices that vary over time but not significantly over across counties. Other important factors vary over both counties and times such as lagged crop acreage. Finally, unobservable county-specific constant effects are important to control for. These considerations lead to the following panel econometric model:

$$A_{it} = \alpha_i + \beta Z_{it} + \gamma X_t + \delta A_{i,t-1} + \theta S_{i,t-1} + \varepsilon_{it} \tag{4}$$

where Z_{it}, as before, is effective, ethanol capacity relevant to county i in year t like in equation (3), X_t is a vector of covariates that are constant across counties, $A_{i,t-1}$ is once-lagged corn acreage, and $S_{i,t-1}$ is once-lagged soybeans acreage.

The inclusion of county-level fixed effect in equation (4) implies that the identification of the ethanol–acreage effect does not come from cross-sectional variation in corn acreage (see the discussion of fixed effects models in Angrist and Pischke, 2009, chapter 5). The causal effect is identified from variations in county-specific effective capacity over time inducing variations in county acreage.

To reduce serial correlation in the disturbance term of equation (4), we take first differences:

$$\begin{aligned} \Delta A_{it} = {} & \beta \Delta Z_{it} + \gamma \Delta X_t + \delta \Delta A_{i,t-1} \\ & + \theta \Delta S_{i,t-1} + \Delta \varepsilon_{it} \end{aligned} \tag{5}$$

Finally, we incorporate the effect of ΔX_t, which are constant across counties, in a completely nonparametric way by replacing ΔX_t with a set of time dummy variables. The full panel model that we estimate is:

$$\begin{aligned} \Delta A_{it} = {} & \sum_{t=2002}^{2007} \varphi_t d_t + \beta \sum_{j=1}^{N} max\left(1 - \gamma D_{ij}, 0\right) \Delta C_{jt} \\ & + \delta \Delta A_{it-1} + \theta \Delta S_{it-1} + u_{it} \end{aligned} \tag{6}$$

where:
i = 1, 2. . ., 2193 (counties);
t = 2002, 2003. . ., 2007 (six periods);
ΔA_{it} = differenced planted corn in period t, measured in acres;
d_t = time dummy variables;
ΔC_{jt} = differenced ethanol production capacity of plant j in period t, measured in annual million gallons;
D_{ij} = distance between county i and plant j, measured in miles;
$\Delta A_{it\text{-}1}$ = differenced lagged value of planted corn and soybeans in period t, measured in acres; and
u_{it} = idiosyncratic error for county i in period t.

The panel comprises seven years (2002–2008) and 2193 counties (the number of counties that meet the criterion of producing corn at least one year during 2002–2008).

Our empirical method is to estimate model equation (3) by nonlinear least squares using annual panel data at the county and plant level. Model equation (3) is linear in parameters conditioned on a value for γ. Therefore, we can calculate linear estimates with different radii and compare the fit of the model with the data for each radius. The criterion for choosing the radius will be to minimize the sum of squared errors (SSE) that result from the model.

The sign of the coefficient β is expected to be positive because additional ethanol capacity is expected to increase planted corn. However, the larger the distance between county i and plant j, the smaller the effect on planted corn in county i. Consequently, the sign of the coefficient γ, which is the distance discount factor, is also expected to be positive. The coefficient on lagged corn acreage should be negative to reflect the negative effect of planting corn after corn. It could also capture serial correlation from unspecified sources. The coefficient on lagged soybean acreage, è, should be positive to reflect the rotational benefits from a corn–soybean rotation.

One final econometric consideration is spatial autocorrelation. Because counties are geographic neighbors, they are likely to have unobservable similarities. In any given period, two neighboring counties are likely to have positively correlated error terms. We account for spatial autocorrelation by calculating confidence intervals for the estimators using a block bootstrapping technique (Hall, 1985). This approach resamples observations with replacement from the original data set in a block form. For each observation (county) withdrawal from the data, the spatial block (a county and its neighbors) is pulled out. In this way, the bootstrapping incorporates spatial autocorrelation into the calculated standard errors.

Our empirical method maintains that the parameter β is a causal effect: an exogenous million gallon increase in effective capacity surrounding a county will induce a β increase in planted corn acres in the county. This causal interpretation rests on an assumption that ΔZ_t and u_{it} are contemporarily uncorrelated. This is to say that changes in planted corn in the current period do not result in the building of ethanol plant capacity in the current period. Although it surely is true that ethanol plants are located near corn-producing counties and so in this sense corn causes ethanol, that does not imply that this year's deviation from expected corn acreage attracts ethanol capacity this year. Our assumption that ΔZ_t and u_{it} are uncorrelated is supported by literature on the ethanol industry documenting the several years it takes to plan and build new plants and expand existing plants. Although we find this argument convincing, we also estimate an instrumental variable version of our model to investigate the probability of simultaneous determination of ΔZ_t and u_{it}.

Data

Data on ethanol plants for the years 2002–2008 are obtained from the Renewable Fuel Association (RFA). The data set is built on seven annual reports (every February) for the years 2002–2008 and includes information on 143 plants that produced ethanol during the years of analysis (the full plant list is available from the authors). The variables available are the plant's name, feedstock used in production (we used only plants using corn as a feedstock), and nameplate capacity (the maximum production level the plant is permitted to produce). Several ethanol firms had multiple plants in the data set and only one aggregate name plate capacity for all of them. In those cases, capacities for individual plants were recovered by contacting the firms or by consulting their web sites. The locations of U.S. ethanol plants are based on information from the RFA and Ethanol Producer Magazine. Latitude and longitude coordinates of the plants were collected separately using the plants' addresses. Although we use plant capacity to represent ethanol plant demand for corn, it is true that plants may operate full, under, or over capacity. There was a trend of increasing capacity use in the ethanol industry since 1999 (Koplow, 2006). Capacity use rate increased from 86% in 1999 to 107% in 2005 and 114% in 2008. Capacity use was close to 100% during our 2002–2008 sample.[4]

The most spatially disaggregated corn acreage data are available at the county level, which we obtain for the 48 continental states from the USDA. Relevant counties for the analysis are those that produced corn in at least one of the years during 2002 and 2008; 2193 counties met this criterion. The data set includes county-level planted corn and soybeans in acres for the years 2001–2008 (year 2001 was necessary for lagged values for year 2002).

Data on ethanol plant capacities are reported once a year, usually in February. Corn is grown usually from May to September and harvested between September and October; therefore, only annual data on corn are available. Consequently, the data on ethanol and corn are synchronized with one another for the purpose of the analysis. The information contained in the February report is available to farmers at planting time.

We use the Great Circle method to calculate distances between two coordinates. Calculating the shortest distance between two points on a sphere, $Dist_{ij}$, can be done by using the following formula:

[4] The rate calculated as the total U.S. ethanol production for a calendar year divided by the total nameplate capacity reported by the Renewable Fuel Association on February of the same year.

$$\text{(7)} \quad Dist_{ij} = ER * ACOS\left\{ \begin{array}{c} COS[(90 - Lat_i)] * COS\left[(90 - Lat_j)\right] + \\ SIN[(90 - Lat_i)] * SIN\left[(90 - Lat_j)\right] * COS[(Long_i - Long_i)] \end{array} \right\}$$

where ER is the earth's radius measured in miles.

Empirical Results

From multiple least squares estimates of equation (4), each conditional on γ, the radius found to minimize SSE is 286.17 miles (equivalent to $\gamma = 0.003494$). The results imply that the weight of actual ethanol capacity in effective capacity, which equals one in the county centroid, diminishes linearly to zero as the distance between the county centroid and the ethanol plant approaches 286.17 miles. Table 2 represents the nonlinear least squares (NLS) results based on the SSE-minimizing choice of $1/\gamma =$ 286.17 miles.

The results show a statistically significant and positive effect of ethanol plant capacity on the level of planted corn nearby. According to the estimates, adding an additional million gallons of capacity at the county centroid stimulates another 5.21 acres (with a standard error of 0.68) of planted corn in the county itself. If a typical 100-million gallon dry mill ethanol plant were built at a county's centroid, then the expected increase in the level of planted corn in the same county would be 521 acres.

According to planted corn statistics, a 521-acre change is economically significant, especially for low-planted corn counties. For instance, in 2008, approximately half of the counties analyzed planted corn between zero and 10,000 acres. A change of 521 acres of planted corn predicted to result from building a 100-million gallon ethanol plant would account for a 5.21% increase in planted corn in a county with 10,000 acres of corn. It would also increase corn acreage in surrounding counties, up to these counties whose county centroids are 286 miles from the new plant. The majority of ethanol plants are located in the Midwest where corn is abundant; therefore, the impact on these counties is lower percentage wise. For more details on U.S. planted corn by county, Figure 5 presents histograms for 2008.

The results in Table 2 also indicate that the soybeans coefficient, θ, is positive and highly significant. This implies that farmers tend to rotate corn and soybeans. We also include lagged corn acreage as a regressor to control for rotation effects. In principle, this gives rise to a dynamic response to ethanol capacity with change in corn acreage feeding forward into subsequent years' changes. Because the lagged corn coefficient is both small and statistically insignificant, we do not distinguish here between short-run and long-run effects.[5]

Using the model estimates, we can aggregate across counties to calculate the response of total corn supply to the siting of a 100-million gallon capacity ethanol plant. The calculation proceeds from the following formula:

$$= 5.21 \times 100 \times \left[max\left(\sum_{i=1}^{N} 1 - 0.003494 \times Distance_{ij}, 0 \right) \right]$$

[5] To investigate the robustness of our results to the assumption of the predeterminedness of changes in effective capacity with respect to innovations in corn acreage, we estimated an instrumental variables version of the empirical model. We used a two-stage method in which the first stage comprised a linear regression of plant capacities Cjt on a variable in Appendix 2. The predicted values from this regression were then aggregated to the county-year level by using the estimate of γ that comes from our nonlinear GLS estimates. In the second stage, a linear panel regression (specification 4) is estimated where the Cjt variables are replaced by their predicted values from the first stage. The resulting IV point estimates are quite similar to those reported in Table 2, although they are less precisely estimated. The key coefficient relating plant capacity to corn acreage (β) is estimated to be 5.21 in Table 2; it is estimated to be 4.62 in the IV version. The standard error in the IV version is larger, resulting in a t-statistics of 2.41 (IV) compared with a t-statistic of 7.63 (NLS)."

Table 2. Nonlinear Least Squares (NLS) Results for 286.17-mile Radius ($\gamma = 3.49 \times 10^{-3}$)

Variable	Estimate	Standard Error	t-Stat
Capacity	5.21	0.683	7.63
Lagged corn	–0.011	0.008	–1.29
Lagged soybeans	0.237	0.009	24.44
D2	–200.95	146.15	–1.37
D3	961.70	146.07	6.58
D4	40.60	148.02	0.27
D5	–1471.67	147.64	–9.96
D6	6097.81	153.11	39.82
D7	–4447.42	172.31	–25.81

Dependent variable—planted corn.

where 5.21 is the corn acreage effect on county i resulting from an additional million gallons of capacity at county i's centroid. This effect is multiplied by 100 to represent the change of planted acres of corn by a typical 100-million gallon plant. The term in parentheses discounts the capacity of each county i using the value of 0.003494 per mile of distance and sums the values for all counties. For each county within 286.17 miles from plant j, the effect will be weighted by the distance between county i and plant j using the estimated discount factor $\gamma = 0.003494$.

The following example applies this formula to the calculation of the total corn supply effect of an actual ethanol plant. The plant chosen is operated by Advanced Bioenergy in Fillmore County, Nebraska. It has 100 million gallons of capacity and uses corn as feedstock. The plant appears for the first time on the RFA annual report in 2008. The total number of counties within 286.17 miles of the plant is 348. Figure 6 demonstrates the 286.17-mile radius around the plant's location. The total effect on corn supply according to the formula is 64,623 acres of planted corn in surrounding counties. If each acre is assumed to yield 150 bushels, then this number is equivalent to approximately 9.7 million bushels of corn.

Putting into context the total corn reaction effect of locating a new ethanol plant, we can calculate the corn response as a share of the corn consumed by the plant. Technology enables producers today to produce approximately 2.7 gallons of ethanol from every bushel of corn in a dry mill process; additional coproducts will be produced during the process. To produce a million gallons of ethanol, then, 370,370 (= 1,000,000/2.7) bushels of corn are required. The total corn response effect from the Advanced Bioenergy plant is 64,623 acres, or approximately 9.7 million bushels of corn; to produce 100 million gallons of ethanol, the plant will have to use approximately 37 million bushels of corn. Dividing 9.7 million by the 37

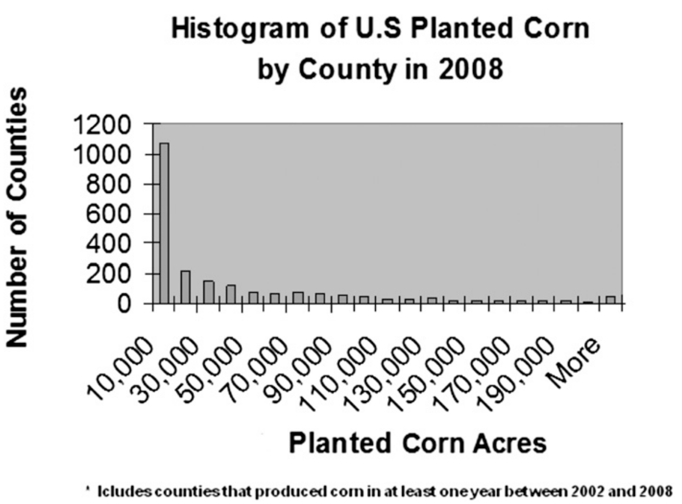


Figure 5. U.S. Planted Corn Histogram (2008)

Figure 6. Radius of 286.17 Miles around the Advanced Bioenergy Plant, Fillmore County, Nebraska: Region of Supply Affected by Plant Siting

million implies that the planted corn reaction effect accounts for approximately 26% of the plant's feedstock needs. In other words, the plant has a corn deficit of 74% of the corn required for production. For the plant to operate at more than 26% capacity, it has to compensate for this corn deficit by attracting corn from other uses or by shipping corn from farther away. The same calculation of corn response as a share of the corn needed for the plant to run at full capacity is done for all ethanol plants in the data set. Figure 7 displays the corn deficit histogram for the plants. Figure 7 indicates that deficit range for all plants falls between 60% and 99%.

This finding is supported by other research, which concludes that ethanol plant siting increases corn prices around the plant (Fatal, 2011b; McNew and Griffith, 2005). Because the corn supply reaction resulting from plant siting does not provide the total amount of corn the plant consumes, the plant makes inroads into competing corn uses, which causes a price increase.

Table 3 presents confidence intervals for the radius and the other estimators from the original model (NLS) using both residual bootstrapping and block bootstrapping. Using the residual bootstrap approach (Anselin, 1990) accounts for estimates' uncertainty by constructing confidence intervals. The computed 90% confidence intervals for the radius (with a point estimate of 286 miles) is 210–389 miles. This confidence interval is constructed by residual bootstrapping of the nonlinear least squares estimates 1000 times and then using the 5% and 95% percentiles of the empirical distribution. See Appendix 1 for details. The bootstrapping procedure is used to calculate the confidence

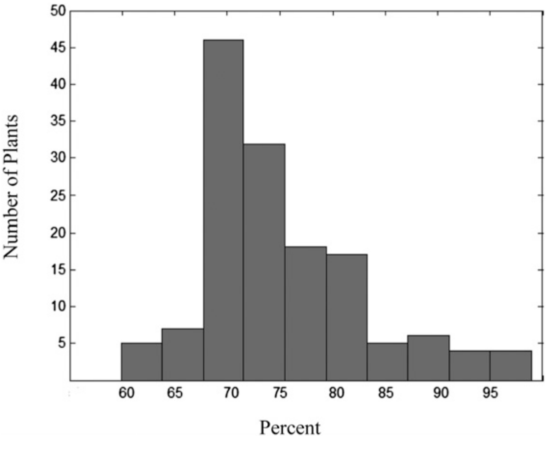


Figure 7. Ethanol Plants' Corn Deficit Histogram: Percentage of Plants Corn Requirements Not Met by Induced Production Increase

Table 3. Confidence Interval Comparison—Nonlinear Least Squares (NLS), Residual Bootstrap, and Block Bootstrap

Variable	NLS CI	Residual Bootstrap CI	Block Bootstrap CI (20-mile neighbors)
Radius (1/γ)	Not Applicable	209.83–389.10	249.99–517.61
Capacity	4.09, 6.34	3.03–7.76	2.4461–10.4220
Lagged corn	–0.0262 to 0.0031	–0.0267 to 0.003	–0.0311 to 0.1709
Lagged soybeans	0.221–0.253	0.2208–0.2535	0.1729–0.2769

CI, confidence interval.

intervals for the coefficients on capacity, lagged corn, and lagged soybeans.

The residual bootstrap will account for sampling variance to the extent that the model disturbances are spatially uncorrelated. The block bootstrap procedure allows for spatial correlation among the disturbances. In it, spatial blocks (a county and its neighbors) are pulled out (Hall, 1985). To perform a block bootstrap, one needs to decide on the criterion that defines a county's neighbors. We choose distance between county centroids. Because the specific distance that defines neighborliness is a matter of discretion, we select a range of reasonable distances and report the results. Table 4 shows the confidence intervals of the block bootstrap estimators using different distances to define neighborliness. Figure 8 demonstrates how the confidence interval for the effective capacity radius shrinks when the distance becomes large enough to include all observations in the data set in one resample. The point where the confidence interval collapses to one point is approximately 2,500 miles and the estimates in this case are identical to the NLS estimates. In other words, using a 2500-mile (or beyond) definition of neighbor counties includes all counties from the original data set and therefore the block bootstrap's data set is identical to the dataset used to generate the NLS estimates.

The confidence interval comparison in Tables 3 and 5 suggest that the sampling variance of the NLS estimator is inflated by spatial autocorrelation. Going beyond correcting the NLS standard errors, spatial disturbance correlation can be exploited to improve the efficiency of estimation. The spatial error model uses maximum likelihood together with a weight matrix to derive estimates that account

Table 4. Block Bootstrap Estimator Confidence Intervals (CIs) Using Different Neighbors' Distances

Distance Definition of Neighbors (miles)	CI Radius	CI Capacity	CI Lagged Corn	CI Lagged Soybeans
20	249.99–517.61	2.4461–10.4220	–0.0311 to 0.1709	0.1729–0.2769
50	130.86–517.75	–4.7456 to 9.1235	–0.0326 to 0.2492	0.1504–0.2986
100	132.46–537.58	−7.6956 to 10.4319	–0.0760 to 0.2976	0.1177–0.3099
200	131.53–534.22	−8.4422 to 10.8812	–0.1312 to 0.3310	0.0822–0.3177
500	148.02–523.77	–5.1570 to 9.1459	–0.0981 to 0.3752	0.1497–0.3203
800	164.10–365.46	–5.1784 to 6.3939	–0.0846 to 0.3495	0.1525–0.3113
1000	168.19–336.31	–5.5307 to 5.5938	–0.0975 to 0.3104	0.1664–0.2967
1500	276.11–286.90	4.5178–5.2665	–0.0809 to 0.0536	0.2347–0.2578
2000	279.40–286.88	4.8615–5.2143	–0.0135 to 0.0345	0.2373–0.2468
2500	286.17–286.17	5.2143–5.2143	–0.0116 to –0.0116	0.2373–0.2373
2600	286.17–286.17	5.2143–5.2143	–0.0116 to –0.0116	0.2373–0.2373

Dependent variable—planted corn.

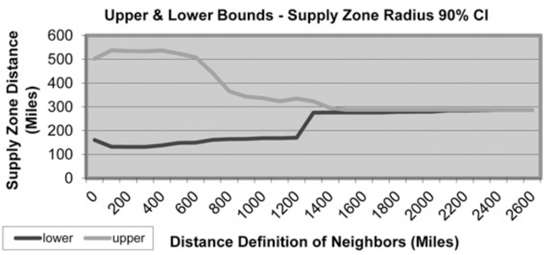


Figure 8. Radius Confidence Interval Values Based on Neighbors Distance Definition

for spatial autocorrelation among counties. We apply two different models using two kinds of weight matrices. The first is based on the inverse distance between counties, and the second (contiguity) acts on binary variables, equal to one if two counties belong to the same state and zero if they do not. The spatial error model uses the radius result from the NLS estimates and calculates new estimates for the coefficients on capacity, lagged corn, and lagged soybeans. The spatial error estimates using the inverse distance and the contiguity weight matrices are similar to the results of the NLS. However, because the spatial error model accounts for spatial autocorrelation among counties, the t-statistics of the coefficient on capacity and lagged soybeans are smaller but still significant at the 1% level (see Table 5).

Discussion

To put our results in context, it is useful to consider industry responses to ethanol demand other than increases in planted corn acreage. Without these other responses, the effect of ethanol industry expansion on planted corn found in this article would have been larger.

First, the amount of corn used for ethanol production can rise not only from increasing supply, but also from corn shifting from other uses. Without shifting corn from other uses, the demand for additional corn resulting from the ethanol industry would have increased corn prices such that it would be more lucrative for farmers to produce corn at the expense of other crops or to allocate more land to corn production. One example of a corn use shift between industries is corn for feed that shifts toward ethanol production.

In practice, the meat production industry is forced to compete with the ethanol industry for corn availability.[6]

Second, corn is diverted to ethanol production not only from other industries. but also from inventories. Even with a corn production increase, the ethanol industry's increased use of corn is reflected in declining ending stocks. Since the establishment of the RFS program and its implementation, U.S. corn ending stocks have decreased from approximately 1.1 billion bushels in 2002–2003 to 880 million bushels in 2010–2011. Furthermore, the ratio between corn stock and total corn use declined from 11.4% in 2002–2003 to 6.6% in 2010–2011. This 6.6% ratio approaches the lowest level of the U.S. corn stock-to-use-ratio of approximately 5% recorded in 1995–1996. Without the opportunity of using additional corn from the ending stock, corn price would increase more and hence lead to a higher effect on planted corn than was observed in our empirical results.

A final point worth mentioning is that the time dummy variables in the model not only control for expected corn price, but also for corn yield changes over time. The total quantity of corn produced in the United States has increased significantly over the last few years. Table 1 shows the increase in production from marketing year 2002–2003 of approximately nine billion bushels to more than 13 billion bushels in 2007–2008. One major factor behind this increase is the rise in corn yield. Table 1 shows an increase in corn yield from approximately 130 bushels per acre to 150 bushels (Kucharik and Ramankutty, 2005). As indicated by the USDA (2007), technological improvements such as new seed varieties, fertilizers, pesticides, and machinery, together with better production practices such as reduced tillage, irrigation, crop rotations, and pest management systems enabled a higher yield. Without including time dummy variables in the model, the analysis would have failed to account for corn yield changes over the years.

[6] It should also be noted that the meat production industry benefits from Dried Distillers' Grains for feed that results as a coproduct in the ethanol production process.

Table 5. Comparison between Nonlinear Least Squares (NLS) and Spatial Error Estimates

Variable	NLS	t-Stat NLS	Spatial Error Model (contiguity)	t-Stat Spatial Error Model (contiguity)	Spatial Error Model (inverse distance)	t-Stat Spatial Error Model (inverse distance)
Capacity	5.2143	7.630128	3.9682	3.5881	3.9582	4.4371
Lagged corn	–0.01156	–1.29686	0.016517	2.4341	0.015926	1.7896
Lagged soybeans	0.2372	24.44493	0.20863	18.642	0.20883	21.218

Dependent variable—planted corn.

As a result of that, the new effect of ethanol plant siting on planted corn would have been smaller than the effect with the time dummies because higher corn yield requires less planted corn for any given production level.

Conclusion

Increasing ethanol production is mandated by the RFS. Unless a different feedstock is used, more corn needs to be either produced or diverted from other uses into the ethanol industry. Today, corn production shows signs of adjustment to meet the regulatory increase in demand. However, with a declining corn stocks-to-use-ratio and increasing ethanol production mandate levels, concerns about the smooth growth of the corn-based ethanol industry have arisen.

The results of this study imply 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. The response to an additional million gallons of capacity at the county centroid is 5.21 acres of corn in that county. However, the effect declines linearly to zero as the distance of the county from the plant approaches 286.17 miles. Moreover, the aggregate effect on planted corn is much higher than 5.21 acres because the supply effect is felt across, typically, hundreds of counties around the plant.

The second RFS mandate requires a production of 15 billion gallons of corn-based ethanol by 2015. It is unclear how the new corn market equilibrium will look when the ethanol production mandate reaches the 15-billion gallon level and should it remain there for following years. Assuming that ethanol producers will indeed expand their capacity, it will be interesting to return to this analysis a few years from now and see how the effect of larger ethanol capacity influences the willingness to plant corn and how sustainable higher levels of planted corn will be over time.

[Received February 2012; Accepted February 2014.]

References

Angrist, J.D., and J.-S. Pischke. *Mostly Harmless Econometrics*. Princeton, NJ: Princeton University Press, 2009.

Anselin, L. “Some Robust Approaches to Testing and Estimation in Spatial Econometrics.” *Regional Science and Urban Economics* 20,2(1990): 141–63.

Urbanchuk, J.M., and J. Kapell. *Ethanol and The Local Community*. Mount Laurel, NJ: AUS Consultants and SJH & Company. June 20, 2002.

Fatal, Y.S. 2011a. “The Ethanol Industry—Location Entry and Plant Capacity Models.” PhD dissertation, Chapter 4, North Carolina State University.

———. 2011b. “The Effect of Ethanol Plant Siting on Corn Basis.” PhD dissertation, Chapter 3, North Carolina State University.

Hall, P. “Resampling a Coverage Pattern.” *Stochastic Processes and Their Applications* 20,2(1985):231–46.

Koplow, D. “*Biofuels—At What Cost*?” Government Support for Ethanol and Biodiesel in the United States. Prepared for: The Global Subsidies Initiative (GSI) of the International Institute for Sustainable Development (IISD) Geneva, Switzerland, October 2006.

Kucharik, J.C., and N. Ramankutty. “Trends and Variability in U.S. Corn Yields Over the Twentieth Century.” *Earth Interactions* 9(2005): 1–29.

Lambert, D.M., M. Wilcox, A. English, and L. Stewart. “Ethanol Plant Location Determinants and County Comparative Advantage.” *Journal of Agricultural and Applied Economics* 40,1(2008): 117–35.

McNew, K., and D. Griffith. "Measuring the Impact of Ethanol Plants on Local Grain Prices." *Review of Agricultural Economics* 27,2(2005):164–80.

Shapouri, H., G. Paul, and G. Michael. *Ethanol Cost-of-Production Survey*. U.S. Department of Agriculture, Agriculture Economic Report Number 808, 1998.

The Clean Fuels Development Coalition and The Nebraska Ethanol Board with cooperation of the USDA, "*Guide for Evaluating the Requirements of Ethanol Plants,*" Summer 2006.

U.S. Department of Agriculture, Ethanol Transportation Backgrounder, Expansion of U.S. Corn-based Ethanol from the Agricultural Transportation Perspective. Internet site: http://www.ams.usda.gov/AMSv1.0/getfile?dDocName=STELPRDC5063605. Accessed November 14, 2009).

U.S. Department of Agriculture, Economic Research Service. 2007. Briefing Rooms Corn Background. Internet site: http://www.ers.usda.gov/topics/crops/corn/background.aspx (Acessed September 10, 2010).

Appendix 1. Recursively Bootstrapping the Dynamic Panel Model

Because it is not possible to calculate confidence intervals or standard errors of either γ or the radius from the conditional least squares methodology described in the text, we developed a complementary bootstrapping methodology. Our approach is one of resampling with replacement of the residuals from the estimated nonlinear least squares model. We then construct the dependent variable recursively using all other explanatory variables and estimates from the nonlinear least squares results.

Let $e_t = [e_{1t}\, e_{2t} \ldots e_{nt}]'$ be the (nX1) residual vector from the model, where t = 2, 3 …T, and

e_t^r be the resampled residual with replacement version of the vector e_t.

The resampled dependent variable (Y) can be constructed recursively using the following relations:

$$Y_2^r = X_2\,\beta_1 + \beta_2 Y_1 + e_2^r$$

$$Y_3^r = X_3\,\beta_1 + \beta_2 Y_2^r + e_3^r$$

$$Y_4^r = X_4\,\beta_1 + \beta_2 Y_3^r + e_4^r$$

$$Y_5^r = X_5\,\beta_1 + \beta_2 Y_4^r + e_5^r$$

$$Y_6^r = X_6\,\beta_1 + \beta_2 Y_5^r + e_6^r$$

$$Y_7^r = X_7\,\beta_1 + \beta_2 Y_6^r + e_7^r$$

and

$$Y = \left[Y_2^r \ldots Y_7^r\right]$$

The vector Y_1 is taken from the original data set and $Y_2^r, \ldots, Y_6^r$ are generated recursively, X is the matrix of explanatory variables except for lagged corn in year t, β_1 is the nonlinear least squares estimator (excluding lagged corn) from the actual data, and β_2 is the nonlinear least squares estimator for lagged corn. Nonlinear least squares is applied to the resampled data. The algorithm is repeated many times to generate bootstrapped estimates, from which are derived the confidence intervals reported in Table 3.

Appendix 2. Instrumental Variables Version of the Empirical Model—List of Variables Used

Variable	Definition
Plants' Competition_it	Indicates whether an ethanol plant already exists in county i and year t (yes = 1, no = 0)
Cattle_it	Number of cattle in county i and year t
MTBE Ban_it	MTBE ban adoption in county i and year t, binary variable (yes = 1, no = 0)
Excise Tax Credit_it	Ethanol producers excise tax credit in county i and year t (yes = 1, no = 0)
Producer Tax Credit_it	Ethanol producer credit program in county i and year t (yes = 1, no = 0)
Wage_it	Annual average wage in county i and year t, measured in U.S. dollars
Per Capita Income_it	Per-capita income in county i and year t, measured in U.S. dollars
Agriculture Region_i	Binary agriculture region indicators of county i and year t. The nine regions are Heartland, Northern Crescent, Northern Great Plains, Prairie Gateway, Eastern Uplands, Southern Seaboard, Fruitful Rim, Basin and Range, and Mississippi Portal; the analysis uses the Heartland region (Reg1) as a base for the rest of the regions