Spatial Dimensions of US Crop Selection:
Recent Responses to Markets and Policy[1]

Mesbah Motamed
Research Economist,
Market and Trade Economics Division,
Economic Research Service, USDA
mmotamed@ers.usda.gov

Lihong Lu McPhail
Research Economist,
Market and Trade Economics Division,
Economic Research Service, USDA
lmcphail@ers.usda.gov

Selected paper prepared for presentation at the Agricultural & Applied Economics Association's 2011 AAEA & NAREA Joint Annual Meeting, Pittsburgh, Pennsylvania, July 24-26, 2011.

[1] The authors gratefully acknowledge the assistance of Ryan Williams of the Economic Research Service, USDA for providing invaluable data and GIS support.

Abstract: We explicitly measure corn acreage response to the biofuels boom from 2006 to 2010. Specifically, we use newly available micro-scale planting data over time to test whether corn cultivation intensifies in proportion to the proximity of ethanol processors. We control for the endogeneity of plant location to corn acreage by using transportation network data for instruments. Our results show that reducing the distance between a farm and an ethanol plant by one percent increases acreage in corn by 0.64% and reveal a price elasticity of supply of 0.47%. To our knowledge, this is the first study that measures changes in location and intensity of corn planting in response to incentives posed by the recent biofuels boom. The results can serve as a springboard for researchers and policy-makers concerned with crop diversity, environmental sustainability, and greenhouse gas emissions.

Key words: corn acreage, ethanol, panel data analysis, instrumental variables

JEL code: Q1, Q28, C33

Introduction

This paper examines the spatial dimensions of U.S. crop selection in response to recent events in energy markets and US farm and energy policy, with a focus on the role of ethanol plants. Specifically, we train our attention on micro-scale outcomes, combining satellite images of crops with observations on ethanol plant locations to measure how crop selection has changed over space and time. We estimate the impact of proximity to ethanol plants on the crop selection decision of producers, accounting for the endogeneity of plant location by using road and water transport networks as instruments and controlling for distances to elevators, prices, ethanol plant capacities, and geographic characteristics.

To motivate this research visually, Figure 1 shows the change in the United States Midwest's corn acreage over the period 2006 to 2010 at the level of individual grid cells of dimension 10 x 10 kilometers. Overall, about 8.2 million acres in corn were added in this region, with the average cell increasing its corn planting area by 410 acres. But as the map shows, the distribution of acreage changes is not uniform; some locations, depicted in dark brown, gained acreage in corn, such as northern Illinois and eastern Iowa, while other spots, colored in lighter shades, saw their area fall, including southern Minnesota and eastern Nebraska. Figure 1 also depicts the location of ethanol plants over the same time period as well as their refining capacities. The light green bar represents a location's capacity beginning in 2006. The second bar indicates that location's capacity at the end of the time period.

From the map, it appears that changes in corn acreage spatially correlate with the presence, introduction and capacity of ethanol plants. Probably the entire reason for this is transport costs. Moving feedstock from the farm gate to the refinery is costly, and all other things being equal, minimizing the distance feedstock must move is part of the optimization problem solved by both crop farmers and ethanol producers.

The intuition underlying our hypothesis tests is thus straightforward. From the perspective of a profit-maximizing crop farmer, proximity to a market, in the form of an ethanol plant and the lower transport costs associated with it, may be an important criterion for choosing to grow corn. In this paper, we measure how much corn acreage responded to the introduction and capacity expansion of an ethanol plant. Our results show that acreage in corn not only gravitated towards locations that introduced and expanded ethanol capacity, but that rotations in areas where corn is already planted actually intensified.

Background

To date, research on crop response to biofuels policies has focused on simulating the market-mediated effects on overall production and deriving estimates of greenhouse gas impacts attributable to the production changes. Given the global nature of the question and the interaction between agriculture and other industry sectors, partial and computable general equilibrium models have been utilized to address these questions including Searchinger, Heimlich et al.

(2008), Keeney and Hertel (2009), Deschenes and Greenstone (2007), and Schlenker, Hanemann et al. (2006). The emphasis here has been identifying and measuring the indirect effects of planting decisions on global environmental outcomes. Analytical presentations have also illustrated producer responses to market and policy shocks (Feng and Babcock 2010). While this research shines considerable light on the nature of producer responses, efforts to empirically capture producers' response remains constrained by data availability and quality (Birur, Hertel et al. 2008). Moreover, the impact on the spatial distribution of planting outcomes in response to biofuels policy remains thus far completely unexamined.

More generally, however, a broader strand of research has focused on measuring acreage response to price and policy variables. Tegene, et al. (2003) derive a model of optimal dynamic agricultural supply assuming farmers have two annual stochastic crop production activities, a joint limitation on production capacity, interdependencies between past acreage utilization and current productivity, and rational expectations. As in many papers, the authors derived a theoretical model of an individual farmers' decision on crop acreage, but relied on state-level acreage data, in this case covering Iowa from 1948-80, to estimate response functions. Chavas and Holt (1990) estimated a system of risk-responsive acreage equations for corn and soybeans in US, paying particular attention to the truncation effects of government price supports on the distribution of corn and soybean prices. A parallel strand of research studies agricultural shares (e.g. corn versus soybeans) as a function of land rents from alternative uses, input and output prices, policy variables, and land quality measures (White and Fleming 1980; Lichtenberg 1989; Stavins and Jaffe 1990; Wu and Brorsen 1995; Wu and Segerson 1995; Plantinga 1996; Miller and Plantinga 1999).

Data

Crop Selection

To capture producers' crop selection, we used the National Agricultural Statistics Service's annual Cropland Data Layer which reports crop location and type at a resolution of 30-square meters across. These data are sensed remotely by satellites, classified into crop types according to multi-spectral rules, and ultimately ground-checked for validity. The highest quality data cover the most agriculturally intensive areas of the United States, namely the Corn Belt and the Mississippi River Delta. Overall, the spatial coverage of the satellite data varies with the years, but a consistent time series of plantings from 2006 to 2010 exists for twelve states that span the Corn Belt. This period coincides with the boom of the ethanol industry and thus allows us to capture the year-to-year response of producers to the (1) introduction of an ethanol plant and (2) and an ethanol plant's capacity expansion.

The 30-square meter observations were aggregated to 10 x 10 kilometer (=100 square kilometer, or about 39 square miles) grid cells. Using high-resolution, regular spatial units such as grid cells offers several advantages. First, we can observe movement and concentration of crop selection within counties, a valuable feature particularly as counties grow larger towards the western half of the study area, and consequently obscure more variation. The average county size in Iowa, by way of illustration, is about 570 square miles, implying about 14 grid cells fall in a typical county. Equally important is the regularity of the spatial unit. Grid cells are arbitrarily-defined, consistent units of observation. They do not reflect an individual farm's size or the political boundaries of a county or state, and are thus independent of any farm or county-level characteristics. They are not defined by any natural features of geography or climate, *e.g.* rivers, elevation, which might also bias crop selection. By virtue of these units, our observations will not be biased by some of the competing explanations for crop selection, such as farm size, management, soil type, or administrative-level policy determinants.

The variable of interest here is the acreage of corn planted in each 10 x 10 km cell. Figure 2 presents a map of corn acreage planted in 2010 and reveals the Corn Belt, beginning in Ohio and stretching westward to Nebraska, South Dakota, and Minnesota. In 2010, about 3,400 corn acres were planted on an average 10 x 10 km cell. The highest valued cell, reporting 18,248 acres planted in corn, appeared in DeKalb County, Illinois. See Table **1** for a summary of the data. Of all the states in the study region, Iowa had cells with the highest average area planted in corn, just over 9,000 acres.

Ethanol Plant Location and Capacity

We use data on ethanol plant location, online date and their year-to-year capacities, reported by the Renewable Fuel Association, to calculate the distance between each cell and the nearest plant and its capacity (Breneman and Nulph 2010). The distance between a producer and the nearest ethanol plant captures the ethanol-driven economic incentives posed to corn growers as mediated by transportation costs. We also include the annual production capacity of the plant to capture the "size" of the market, insomuch as larger capacity plants probably exert influence over a wider radius of farms.

In 2006, 84 ethanol plants appeared in the data set, with a total nameplate capacity of about 4.5 billion gallons per year. By the end of 2010, the number of plants rose to 158 with a corresponding nameplate capacity of 11.4 billion gallons. From 2006 to 2010, the average distance between a grid cell and its nearest ethanol plant fell from 138 kilometers to 96 kilometers, and the average total capacity of the nearest grid cell location that hosted plants rose from 41 million to 59.6 million gallons per year.

In 2010, Iowa led the region in ethanol plants, hosting 39, with the average plant-hosting grid cell capacity of 85 million gallons per year. In Iowa, the average grid cell is just 6 kilometers away from the nearest ethanol plant. A yearly summary of the plants and their capacities for the whole sample as well as state-level numbers appears in Table 2.

Endogeneity of Ethanol Plant Location and Selection of Instruments

Ethanol plants do not select their location independently of their surroundings (Breneman and Nulph 2010). Certain features of a location, namely its transport infrastructure, feedstock supply, water availability, policy incentives, and of course, the presence of other plants, drive investors' decisions to introduce a new plant in a region or upgrade its capacity. Stewart and Lambert (2011) show that an area's corn production positively influences a plant's decision to locate, and as such, our left-hand-side variable, planted acreage in corn, introduces a possible endogeneity bias into our estimates.

To get around this issue, we introduce a set of instruments that plausibly correlate with plant location decisions, namely transportation infrastructure, that are at once uncorrelated with producers' planting decisions. As detailed in Stewart and Lambert (2011), road density, and rail and river networks, among other transport-related attributes, are important determinants of plant location. For the instrument to be valid, it must be the case that planting decisions (corn versus other crops) are uncorrelated with the local transportation infrastructure. Since planting decisions are generally a function of expected price and agronomic characteristics, the role of proximate transportation infrastructure in determining a particular crop's selection, as opposed to, say, the decision to enter or exit agriculture, is believed to be small. In fact, the correlations between corn acreage and the distance to the interstate highways, secondary intersections, and water ports are -0.21, -0.21, and -0.23, respectively.[2]

Moreover, given the relatively time-invariant characteristics of these transport infrastructures, and the fact that nearly all highways and rivers pre-dated the construction of ethanol plants, the dependence of transportation network on ethanol plant locations is safely assumed to be zero. This permits us to estimate a causal relationship between ethanol plant location and transportation infrastructure in the first stage of the estimation.

Values for each cell's distance to the nearest interstate, secondary intersection, and water ports were constructed based on data layers provided by ESRI's ArcGIS software package.

Additional Controls

To control for additional factors that explain crop selection, we include variables that report the distance from each 10 x 10 km cell to the nearest price point, usually an elevator, and the cash grains bid recorded at the point. Price data are from Cash Grain Bids, Inc.[3] Soil quality data are taken from the National Commodity Crop Productivity Index to control for time-invariant geographic features that capture the agronomic component of corn production. The index ranges from 0 to 1000, with higher values reflecting higher productivity.

Exploratory Analysis of the Data

Tests of Spatial Autocorrelation

[2] When data for rail networks become available, these will be included too.
[3] www.cashgrainbids.com

Since this paper's goal is to report the year-to-year change in spatial concentration of corn selection in response to ethanol plant proximity, the first step is to document whether spatial concentration is actually occurring over time. The most common statistic for this is the Moran's *I*, which essentially captures the correlation of like-values in a given neighborhood (Stewart and Lambert 2011). We calculate the Moran's *I* statistic for corn acreage across the region of interest over the period 2006-2010.[4] See Table **3**. From the results, we can see that corn was already highly concentrated in the region, with a Moran's *I* of 0.65, but that the concentrations dropped and then returned to its original level by the time period's end. Year 2007 witnessed record plantings in corn, a fact which may be reflected in the slightly reduced concentration, as a result of additional extensive planting of corn.

Relating Acreage to Proximity and Capacity

Does the spatial concentration of corn plantings correlate with the proximity and capacity of an ethanol plant? Figure 3 plots a local polynomial curve relating each cell's corn acreage to its nearest plant for each year. From the figure, it appears not only that acreage decreases in distance to the nearest plant, but that each year sees the relationship grow increasingly steep. Fitting a simple line to the data confirms that the average distance between a cell to the nearest plant and the area of corn planted on it strengthens over time, suggesting that acres are spatially concentrating within the vicinity of plants over the time period.

Another way to measure this is at the plant level. How has the distribution of corn acreage varied in relation to a given plant's capacity? To answer this, we first defined a circular neighborhood surrounding each of the 168 grid cells that hosted some plant capacity. The circle's radius was set to 138 kilometers, following the 2006 average distance between the total sample's grid cells and their nearest plant.[5] This equates to about 13 cells in each cardinal direction from the origin cell. We then add up all the acres planted in corn inside each of the cells that fall within the radius of the origin cell. Figure 4 plots this neighborhood sum against the capacity of the origin cell's capacity for each. Not surprisingly, acreage in corn increases in capacity, with the relationship between the two variables apparently strengthening over time.

Estimating an Acreage Response Model

Specification

We construct a panel composed of five annual observations from years 2006-2010 reporting each grid cell's (1) corn acreage; (2) distance to nearest ethanol plant; (3) nearest plant's capacity; (4) distance to nearest elevator; (5) cash grain bids at the nearest elevator; (6) county and state identifiers.

---

[4] We construct a spatial weights matrix using queen contiguity criterion with 5 orders of contiguity to create a neighborhood of 121 grid cells, corresponding to an area of approximately 360 square kilometers.
[5] Any arbitrary number would serve this purpose. We select the 2006 average distance just for the sake of convenience.

To build our instrument, we calculate the distance from each 10 x 10 km cell to each of the following transportation features: (1) nearest interstate ramp; (2) nearest intersection of secondary and primary roads; and (3) water ports. We estimate these three variables' effect on the distance of each 10 x 10 km cell's distance to the nearest ethanol plant. The first stage estimation, thus, appears as follows:

$$ethanol_i = \beta_1 interstate_i + \beta_2 intersection_i + \beta_3 waterport_i + \varepsilon_i$$

Where *ethanol* represents the log of the distance between each 10 x 10 km cell and the nearest ethanol plant, and *interstate*, *intersection*, and *waterport* represent the log of the distances between each cell and the nearest interstate ramp, secondary and primary road intersection, and water port, respectively. The predicted values of *ethanol* then enter into the second stage estimation as $\widehat{ethanol}$. The results for the first stage regression are reported in Table 2. Because the instruments are time-invariant, we use a random effects model for the second stage analysis. The second stage specification, then, appears as:

$$acreage_{it} = \beta_1 ethanol_{it} + +\beta_2 price_{it} + \beta_3 cap_{it} + \beta_4 elev_{it} + \beta_5 nccp_i + \delta_i + \varepsilon_{it}$$

Where *acreage* is the log of corn acreage planted in grid cell *i* in year *t, price* is the log of the ratio of corn to soybean cash bid prices at the nearest elevator for grid cell *i* in year *t*, *cap* is the log of capacity of the nearest ethanol plant for grid cell *i* in year *t, elev* is the log of the distance to the nearest corn elevator for grid cell *i* in year *t*, *nccp* is the average overall soil productivity index for grid cell *i*, δ is unobserved individual grid cell heterogeneity which is assumed to be uncorrelated with the included variables.

Results from the first stage regression to obtain predicted values for ethanol plant distance to transportation infrastructure appear in Table 4. Results from the second stage panel regressions are reported in Table 5.

Results and Discussions

Table 5 reports the panel regression results that begin with the simplest specification and gradually introduce new controls, building to a final specification in Column 5. All estimates emerge significantly at the 1% level with all the expected signs. The results from Column 5, our most comprehensive model, show that shortening the distance between an average grid cell and its nearest ethanol plant by one percent increases the acreage planted in corn by 0.63%. A one-percent rise in the refining capacity at the nearest ethanol plant lifts acreage planted in corn by 0.125%. The remaining estimates confirm the relationship between corn acreage and the corn-bean price ratio and the distance to the nearest elevator. A new result also emerges from the estimate on the corn-bean price ratio. The elasticity, 0.474, represents the average percent effect of a one-percent rise in the relative price of corn to soybeans at the nearest price point to a given cell. This is the first estimate of supply elasticity on this scale and level of observations.

Conclusion

Understanding the effects of the biofuels industry on the landscape of US agriculture remains a priority both for policy makers and researchers. The aggregate response of US producers to the incentives posed by higher ethanol prices and the recently-passed biofuels mandate is unmistakable, but the location of these responses has yet to be fully documented. In this paper, we ask how producers have responded to the presence, introduction, and expansion of an ethanol plant in their vicinity. Controlling for the endogeneity of an ethanol plant's location, we show that acres in corn rose in response to the proximity of an ethanol plant and its capacity. This points to a trend in the spatial concentration of corn acreage in the US. Coupled with recent trends in co-locating livestock feeding operations, the implications for this adjustment in the spatial pattern of planting include more intensive land use in areas surrounding ethanol plants, more concentrated environmental impacts, and tighter linkages between the food, feed, and energy sectors. These outcomes may reflect the efficient response of different producers to new economic incentives, but any externalities associated with these evolving arrangements remain unknown. This paper highlights these changes, relying on annual micro-scale satellite data that facilitate our understanding of planting decision variation over time and space. Further refinements in the analysis will introduce dynamic responses in the year-to-year changes as well as more spatially-explicit analyses that focus on neighborhood-level outcomes.

References

Birur, D., T. Hertel, et al. (2008). "Impact of Biofuel Production on World Agricultural Markets: A Computable General Equilibrium Analysis."

Breneman, V. and D. Nulph (2010). "Analysis of Renewable Fuels Association Ethanol Plant Data." Economic Research Service.

Chavas, J.-P. and M. T. Holt (1990). "Acreage Decisions under Risk: The Case of Corn and Soybeans." American Journal of Agricultural Economics **72**(3): 529-538.

Deschenes, O. and M. Greenstone (2007). "The Economic Impacts of Climate Change: Evidence from Agricultural Output and Random Fluctuations in Weather." American Economic Review **97**(1): 354-385.

Feng, H. and B. A. Babcock (2010). "Impacts of Ethanol on Planted Acreage in Market Equilibrium." American Journal of Agricultural Economics **92**(3): 789-802.

Keeney, R. and T. W. Hertel (2009). "The Indirect Land Use Impacts of United States Biofuel Policies: The Importance of Acreage, Yield, and Bilateral Trade Responses." American Journal of Agricultural Economics **91**(4): 895-909.

Lichtenberg, E. (1989). "Land Quality, Irrigation Development, and Cropping Patterns in the Northern High Plains." American Journal of Agricultural Economics **71** 187-194.

Miller, D. J. and A. J. Plantinga (1999). "Modeling Land Use Decisions with Aggregate Data." American Journal of Agricultural Economics **81**(1): 180-194.

Plantinga, A. (1996). "The Effect of Agricultural Policies on Land Use and Environmental Quality." American Journal of Agricultural Economics **78**: 1082-1091.

Schlenker, W., W. M. Hanemann, et al. (2006). "The Impact of Global Warming on U.S. Agriculture: An Econometric Analysis of Optimal Growing Conditions." Review of Economics and Statistics **88**(1): 113-125.

Searchinger, T., R. Heimlich, et al. (2008). Use of U.S. Croplands for Biofuels Increases Greenhouse Gases Through Emissions from Land-Use Change, Iowa State University, Department of Economics, Staff General Research Papers.

Stavins, R. and A. Jaffe (1990). "Unintended Impacts of Public Investment on Private Decisions: The Depletion of Forest Wetlands." American Economic Review **80**: 337-352.

Stewart, L. and D. M. Lambert (2011). "Spatial heterogeneity of factors determining ethanol production site selection in the U.S., 2000–2007 " Biomass and Bioenergy.

Tegene, A., W. Huffman, et al. (2003). An Econometric Model of Dynamic and Stochastic Production with Explicit Optimization: Farmers' Decisions on Corn and Soybean Acreage, Iowa State University, Department of Economics, Staff General Research Papers.

White, F. C. and F. N. Fleming (1980). "An Analysis of Competing Agricultural Land Uses." Southern Journal of Agricultural Economics **12**(2): 99-103.

Wu, J. and B. W. Brorsen (1995). "The Impact of Government Programs and Land Characteristics on Cropping Patterns." Canadian Journal of Agricultural Economics **43**: 87-104.

Wu, J. and K. Segerson (1995). "The Impact of Policies and Land Characteristics on Potential Groundwater Pollution in Wisconsin." American Journal of Agricultural Economics **77**: 1033-1047.

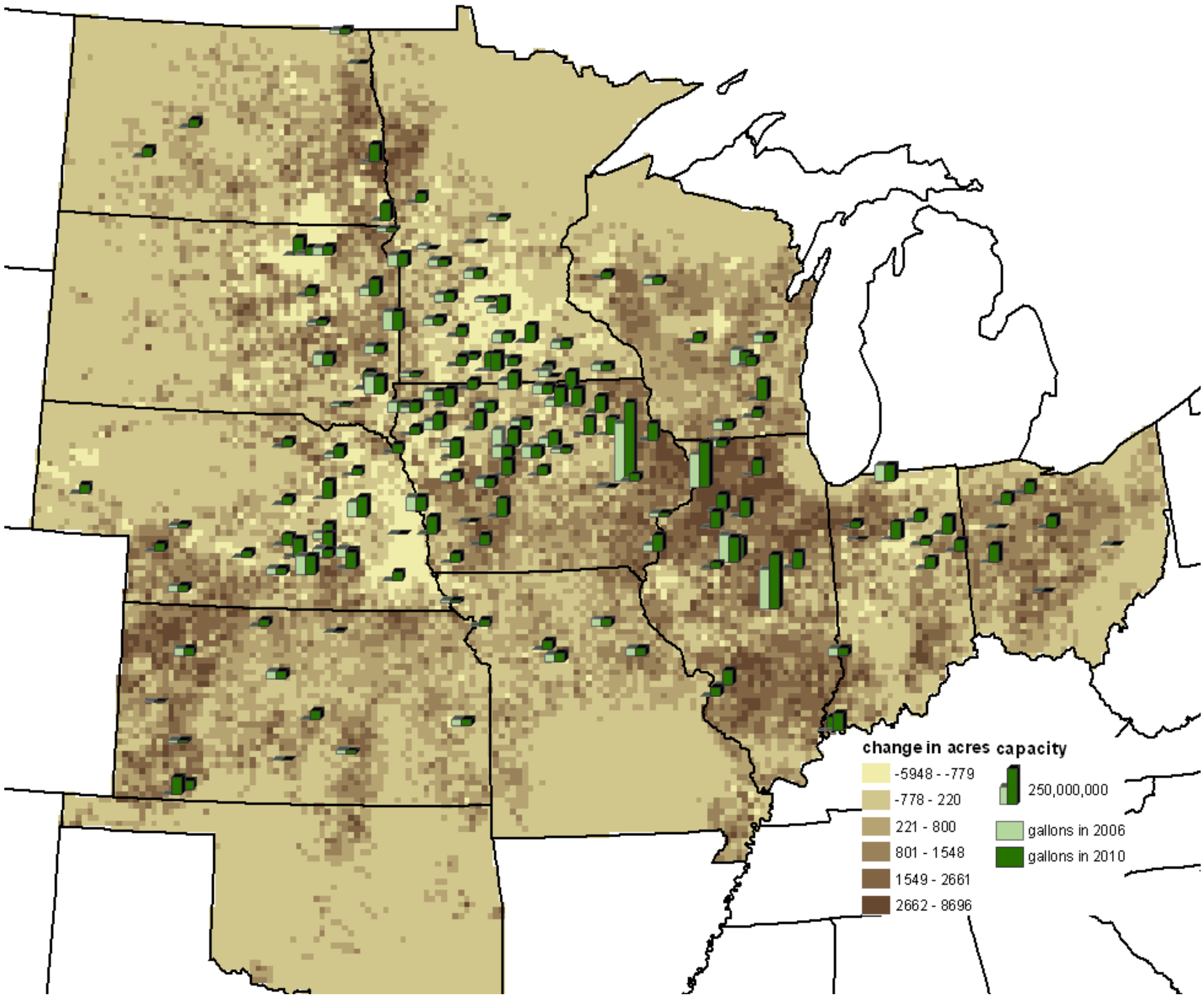

Figure 1. Change in corn acreage and plant capacity

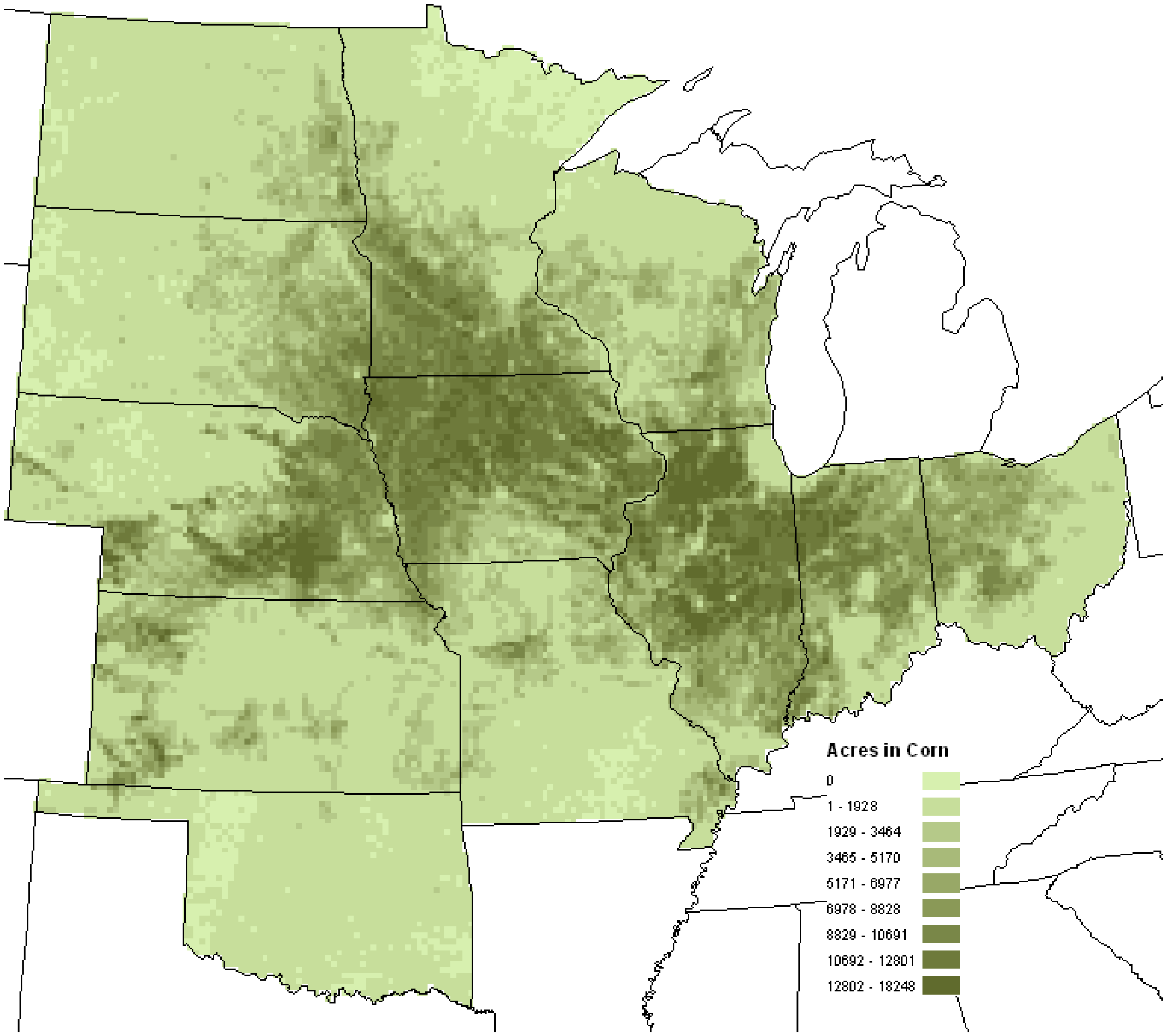

Figure 2. 2010 Acreage in Corn, US Midwest

Table 1. Summary Statistics of Corn Area in Acres

Year	2006	2007	2008	2009	2010
All grid cells[6]					
Mean	2,990	3,326	3,102	3,150	3,399
Maximum Value[7]	18,109	18,896	19,023	1,796	18,248
Standard Deviation	3,813	3,326	3,874	3,834	4,014
State-level means					
Illinois	6,919	7,969	7,473	7,549	8,346
Indiana	5,552	6,152	5,528	5,542	6,109
Iowa	8,349	8,722	8,484	8,612	9,086
Kansas	1,387	1,595	1,748	1,757	2,117
Minnesota	3,403	3,479	3,223	3,095	3,383
Missouri	1,311	1,477	1,152	1,537	1,578
Nebraska	4,322	4,175	4,243	4,284	4,416
North Dakota	728	1,308	1,192	1,112	1,011
Ohio	2,458	3,229	2,858	2,556	3,120
Oklahoma	66	139	122	159	181
South Dakota	2,131	2,556	2,231	2,236	2,296
Wisconsin	2,134	2,599	2,187	2,349	2,587

[6] Total sample size was 20,137 cells.
[7] Minimum values for all states were zero except Iowa, whose minimum-value cell had 12 acres.

Table 2. Annual Plant Numbers and Capacity[8]

Year	2006		2007		2008		2009		2010	
Total sample	84	4,532	97	5,649	128	7,723	165	10,954	158	11,476
State level										
Illinois	9	872	10	979	10	1,058	16	1,541	16	1,730
Indiana	1	102	3	210	8	491	10	679	10	807
Iowa	19	1,235	24	1,837	28	2,152	38	3,261	36	3,326
Kansas	6	167	7	207	10	427	12	487	11	434
Minnesota	16	555	16	561	18	768	22	1,051	21	1,122
Missouri	3	141	3	141	4	166	5	241	5	241
Nebraska	12	571	13	683	19	1,008	25	1,476	25	1,594
North Dakota	1	26	3	126	4	167	6	271	5	353
Ohio	0	0	0	0	5	330	6	383	4	314
South Dakota	12	630	12	630	14	699	15	1,016	15	1,016
Wisconsin	5	233	6	275	8	457	10	548	10	538

[8] Left column indicates the number of ethanol plants. Right column indicates total nameplate capacity in millions of gallons.

Table 3. Spatial autocorrelation of corn acreage

year	2006	2007	2008	2009	2010
Moran's *I*	.65	.62	.62	.63	.65

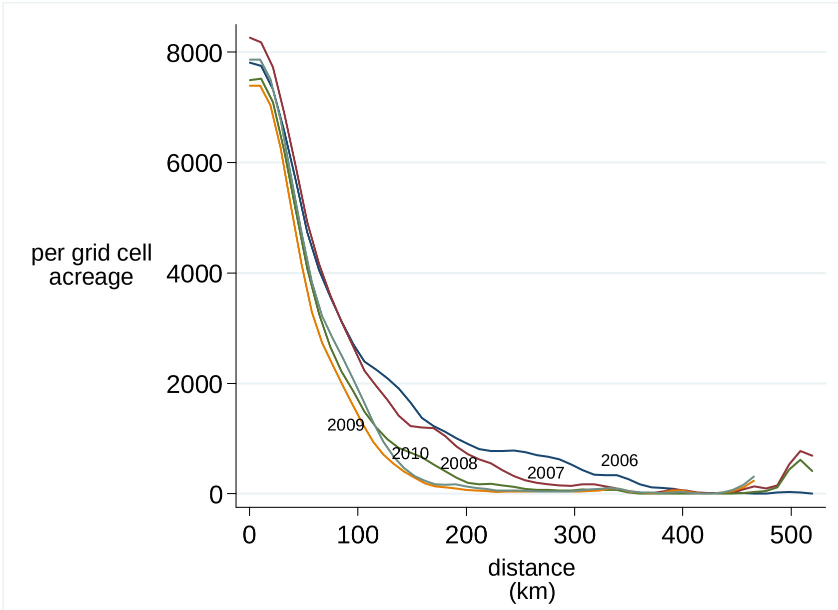

Figure 3. Annual distance from grid cell to nearest ethanol plant

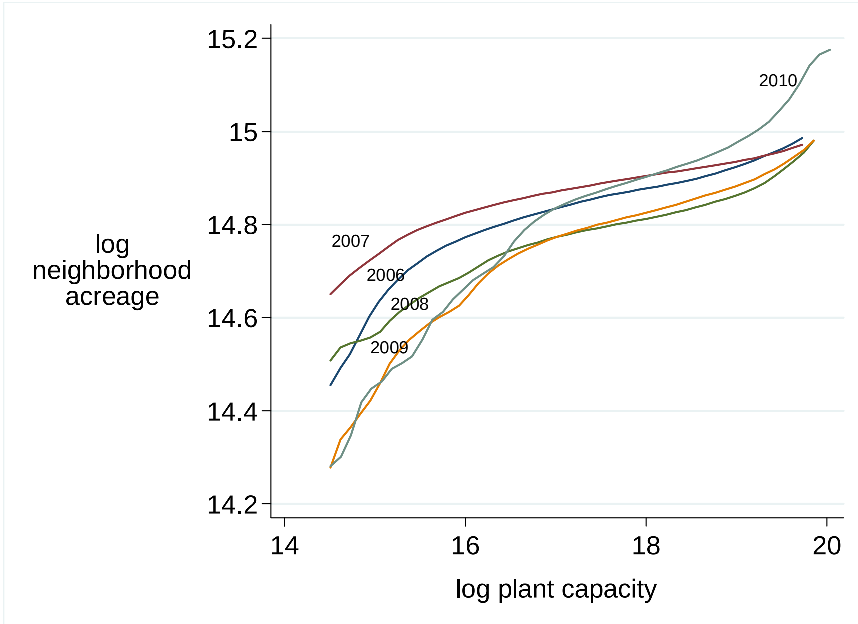

Figure 4. Annual concentration of corn acreage in plant neighborhoods

Table 4. First Stage Regression Results[9]

Variable	Estimate	p-value
interstate	0.025626	0
intersection	0.09624	0
waterport	0.042254	0
constant	9.52289	0
adjusted R-squared	0.02	

[9] The dependent variable is log distance to the nearest ethanol plant.

Table 5. Panel Data Analysis Results

Model specification	1	2	3	4	5
constant	14.966	15.563	12.111	20.481	16.084
log distance to ethanol plant	-0.790	-0.801	-0.753	-0.681	-0.637
log corn-bean price ratio at nearest elevator		0.460	0.487	0.512	0.474
log capacity of nearest ethanol plant			0.168	0.164	0.125
log distance to nearest elevator				-0.934	-0.722
log NCCP soil index					0.001
random grid cell effect	yes	yes	yes	yes	yes
instrumental variables	yes	yes	yes	yes	yes
overall R-square	0.3558	0.3345	0.3437	0.495	0.4686

Note: All estimates are significant at 1% level.