Investment in Corn-Ethanol Plants in the Midwestern United States[1]

C.-Y. Cynthia Lin and Karen Thome

Abstract

Most American fuel-ethanol plants use corn as a feedstock, and thus are located in the Midwestern United States, where the majority of the corn in the US is grown. Since biofuels have been touted as a way to enhance profits in rural areas, where grain prices have remained stagnant over time, it is important to determine what factors affect decisions about when and where to invest in building new ethanol plants. In this paper we model the decision to invest in ethanol plants using both a reduced-form and a structural model. We find that competition between plants is enough to deter local investments. We also find that availability of corn is important in determining plant location.

Keywords: ethanol, biofuels, investment timing game, entry

---

[1] University of California at Davis. Lin: cclin@primal.ucdavis.edu. Thome: thome.karen.e@gmail.com. This research was supported by a UC-Davis Chevron Research Grant. All errors are our own.

1. Introduction

Recently the support of biofuel production has been a politically sensitive topic. Several government policies actively promote ethanol production via tax incentives and mandates. Additionally, politicians have pushed for ethanol as an environmentally friendly alternative to imported oil, as well as a way to boost farm profits and improve rural livelihoods. All these factors have coincided with investment in the industry in recent years.

Fuel ethanol has been in use in the United States since the time of the Model T Ford (the original flex-fuel vehicle), and while the United States passed Brazil in ethanol production in 2005, today ethanol is mostly relegated to status as a gasoline additive. The first US ethanol boom began as a result of the oil embargoes in 1973 and 1979. The desire for more energy self-sufficiency, the resulting legislation (in the form of federal income tax credits and blender's credits that continue today), and phase out of leaded gasoline led to the construction of 153 new plants by 1985 (DOE 2008). These plants were tiny by today's standards, with an average capacity of 8 million gallons per year, and by 1991 only 35 were still operational due to poor business judgment and bad engineering (DOE, 2008; Urbanchuck, 2006).

The second US ethanol boom began in the mid-1990s and hit full-stride by the early 2000s. Several factors contributed to this most recent boom. The Clean Air Act of 1990 mandated use of oxygenates in gasoline, of which ethanol is one, and the subsequent phase out and ban of MTBE as additive beginning in the late 1990s further increased demand for ethanol. Additionally the Renewable Fuel Standards of the Energy Policy Act of 2005 mandated ethanol production floors beginning in 2007, which rise to 36 billion gallons per year in 2033. Over this time period, the number of ethanol plants rose from 35 plants in 1991 to 50 in 1999 to 110 in 2007 for a total capacity of 6.5 billion gallons per year.

In addition to the policy and demand-side contributors to recent ethanol boom, this new industry growth has been accompanied by changes in plant management and technology. Most significantly, the average capacity of plants was 50 million gallons per year in 2007, up from 8 million gallons in 1985. In the mid-1990s the industry began shifting to more efficient plants, which use natural gas instead of coal as fuel (DOE, 2008). Ownership is also shifting to streamlined corporate owners with multiple plants. Historically, farmer-owned plants had a large share of the market, though by 2007 only 11% of new capacity was farmer owned, while the largest 5 corporations had 42% of capacity in 2008 (FTC, 2008).

This recent boom, in addition to industry shifts in technology and ownership, beg an analysis of investment decisions. Most ethanol plants use corn as a feedstock, and thus are located in the Midwestern United States, where the majority of the corn in the US is grown. Since biofuels have been touted as a way to enhance profits in rural areas, where grain prices have remained stagnant over time, it is important to determine what factors affect decisions about when and where to invest in building new ethanol plants. We model this decision in the paper using both reduced form and structural models.

There is a related literature on food manufacturing location decisions which begins with a basic model of determinants of manufacturing establishment growth. One example is Goetz's (1997) analysis of the determinants of rural food manufacturing establishment growth. He considers the effects of the following factors: access to output markets, labor force composition and quality, transportation infrastructure, government intervention, and availability of raw materials. In his model location decisions involve a two-step process where regions are chosen for broader consideration, and then choice is narrowed within each region. Goetz in turn estimates the net change in the number of establishments in each state and county as a function of the previous factors using OLS.

There are two related studies that specifically address the location of ethanol plants. Sarmiento and Wilson (2007) use a discrete choice model to analyze agricultural characteristics and spatial dimensions that determine plant location. While Sarmiento and Wilson analyze similar factors as Goetz, this is a cross-sectional study, and they can not address investment in plants, and instead study location determinants. Similarly, Lambert and Wilcox (2008) use a discrete choice model with spatial clustering to look at factors that affect the presence of ethanol plants and proposed plants in a given county, and also isolate clusters that may attract investment. They consider active plants (and new plant announcements) from 2000-2007 as a cross-section, and used location determinants from 2000 or prior to avoid simultaneity.

These studies provide a starting point for our analysis. Both studies find, in particular, that access to corn is important. Both also find that policies and presence of cattle positively influences plant location. Neither study adequately addresses the potential competition between plants in the location decision. While Lambert and Wilcox consider plants in operation from 2000-2007, they treat this sample as a cross-section. The binary dependent variable is 1 if there is a plant in the county and 0 otherwise; multiple plants in the county only enter into the estimation through weighted observations. Thus, it is not possible to study local competitive effects on location in this framework; in addition, Lambert and Wilcox (2008) do not include potential regional competition between plants as an explanatory variable.

Sarmiento and Wilson (2007) also employ a cross-sectional model with a binary dependent variable. They however do spatially lag the dependent variable (as well as the corn availability variable) in order to estimate the competitive effect between plants; they estimate this using a non-linear logit model, allowing the spatial effects to enter non-linearly with distance. They find that competition between plants is negatively correlated with plant presence, and furthermore, that this effect decreases with distance. It is important to note however that this competitive effect only describes the relationship between existing plants; without panel data it is not possible to analyze the effect of competition on entry. We thus improve upon these models by creating a dynamic estimation scenario with panel data, and by directly estimating the effect of covariates on the investment decision itself.

2. Theoretical Model

Our model begins with whether or not there is an investment in an ethanol plant in county i in year t. Investment in an ethanol plant is irreversible and, in each year t, all investment decisions are made simultaneously. I_{it} is an indicator of whether there is investment in a new ethanol plant in county i in year t.

The profit from investing in an ethanol plant in year t is denoted by $\Pi(\bullet)$, which is the expected revenue from the plant minus the expected costs. More specifically, the expected profit depends on exogenous covariates X_{it} that describe the state of the input and output markets, as well as the number of other plants in the county a_{it}.

The exogenous covariates X_{it} describe the expected profits from the sale of ethanol, as well as the potential for profit from the sale of co-products. The co-product market has become more significant as of late due to oversupply of ethanol and the resulting lower prices (Dhuyvetter, Kastens, and Boland, 2005).

There are two types of corn-ethanol plants, dry mills and wet mills. Dry mills use a mechanical method and natural gas to dry the co-product, producing distiller's grains (DDGS, or distillers' dried grains with solubles), an easily transportable animal feed. Wet-milling is making a comeback however as high-fructose corn syrup prices rise (DOE 2008).

X_{it} also includes covariates describing the cost of ethanol production. One important factor is availability and cost of corn, the primary feedstock in our region of focus. McNew and Griffith (2005) find that corn price increases were spatially concentrated around ethanol plants. Corn is the largest variable cost in ethanol production. Other factors that affect expected costs include natural gas price and market access. Finally, state policies can also affect expected profit from a new investment in an ethanol plant.

It is also important to address the potential strategic interactions/competition between plants. Plants locating nearby have the potential to create positive and negative externalities for entering plants. In terms of positive externalities, there could be benefits for a new plant from taking advantage of the transportation or marketing infrastructure or the educated work force already developed by an existing plant; all these factors can affect manufacturing location decisions (see e.g., Goetz, 1997; Lambert and Wilcox, 2008). In terms of negative externalities, plants could compete in both the output and input markets. For example, Sarmiento and Wilson (2007) explain that their estimated negative competitive effect is due to competition in feedstock procurement.

There will be a new investment in an ethanol plant in county *i* in year *t* if the profits from the investment are greater than a continuation value from waiting. The value function for a county *i* in a given period *t* can be written as

$$V(a_{it}, X_{it}) = \max\left\{\Pi(a_{it}, X_{it}), \beta V^c(a_{it}, X_{it})\right\}.$$

The payoff from building an ethanol plant in county *i* in year *t* is $\Pi(a_{it}, X_{it})$. If there is no investment at time *t*, then the payoff in the county *i* is the discounted continuation value, where β is a discount rate, and $V^c(\bullet)$ is the continuation value to waiting instead of investing this period. Whether or not there is an investment depends on which value is greater in that period.

3. Data

Time-Frame and Focus Region

We are investigating the decision to build and operate fuel ethanol plants. While ethanol is produced in most parts of the United States, and from various feedstocks, 95% of ethanol is produced from corn. By focusing on corn-ethanol plants we eliminate the need to consider feedstock choice in the model. The majority of corn (and ethanol from corn) is produced in the Midwest, so we focus on this region, specifically these ten states: Iowa, Illinois, Indiana, Kansas, Minnesota, Missouri, Nebraska, Ohio, South Dakota, and Wisconsin.

While fuel ethanol has been around since the time of the Model T Ford, a narrower time frame is necessary. The dependent variable is the operational date for corn-ethanol plants, and as such, we focus on plant openings from 1996-2007. This corresponds to the second ethanol boom. This allows us to focus on one set of policy variables as well as on plants with modern technology. Starting the analysis earlier would also be difficult because plant startup and closure information is not readily available before this date.

With the ten states and 13 years of observations, we have a total of 918 counties and 11,934 potential observations. We eliminated completely non-agricultural counties (e.g. Northern Minnesota) as well as those with missing data, leaving 10385 observations in 865 counties.

Plant Data

We constructed a unique panel dataset of information on the investment in ethanol plants ten states in the Upper Midwest (SD, MN, WI, KS, NE, IA, MO, IL, IN, OH) from 1996-2008. These states were chosen based on number of corn-ethanol plants, and volume of corn production. This dataset includes plant start-up date, capacity, start-up date and size of additions, mill type, and ownership type. The original list of operational plants was obtained online from the Renewable Fuels Association and Ethanol Producer magazine, including historical lists from the Renewable Fuels Association; these lists do not match perfectly. We were able to rectify inconsistencies between the two lists as well as collect additional information on plant owners by searching through plant websites, newspaper articles and SEC filings. We have taken care to note not only new plants, but also plant closures.

Our sample begins with 26 operational plants in 1996, and ends with 116 operational plants in 2007 (Figure 1).

We collected similar information on biodiesel plants; the original plant lists were from biodiesel.org and biodieselmagazine.com.

Other Data

We use several variables that might affect the cost of ethanol production. Corn prices and soy prices are available annually from the National Agricultural Statistics Service (NASS) of the USDA at the state level, and natural gas (city gate) prices are available annually from the EIA, also at state level.

While it is unfortunate that these data are not publically available in a more disaggregate form, we can find county variation in the intensity of corn production. Corn production and acreage is available annually by county from the NASS. Since all counties are different sizes, we use intensity to capture area-independent acreage (fraction planted to corn) and production (bushels per county-acre) by county, using county acreage from the US Census. We construct similar measures for soy since soy and corn compete for acreage, and soy can also be used in feed, thus competing with corn and DDGS.

We constructed regional corn and soy intensity variables in the same manner as the county intensity variables. We defined region as all the counties sharing any border with the county, not inclusive of the original county in question.

We also use a sow density variable to represent the potential market for DDGS. This is constructed using the number of cows per county-acre, where the number of cows is from NASS. We would ideally use hog data as well, but this information is not available at the county level for all states.

The ethanol price is the FOB price in Omaha, and is from the Nebraska Energy Office. The oil price is the annual average US crude composite acquisition cost by refiners, from the EIA. Once again, it would be ideal to have more spatially disaggregate prices, but they are not publically available. Since one assumes markets are integrated, our price data at least allows us to study the effect of trends over time. We use the average urban CPI to deflate the prices.

We use the presence of gas and ethanol terminals in a county to proxy for market access. We obtained a list of terminal locations of prices available from the Oil Price Information Service (OPIS).

We considered several potential policy variables. At the federal level, we chose to not include tax breaks, credits or the small-producer subsidy because they do not vary enough in our time period to identify them. We do include a policy variable for the MBTE ban, since the ban came into effect at different times for each state.

The second policy variable we include represents the state producer tax credits. Not all states in the sample have these policies, and those that do were in place for plants that opened in different years. Describing this variable is complicated by the fact that each state places different contingencies on receiving funds; for example, some states support large plants, others only small or community owned plants. Because of these differences we represent these policies with a binary variable.

Summary Statistics

Table 1 presents summary statistics for all relevant variables. There are noticeable differences between counties with and without ethanol plants with respect to the agricultural variables.

Table 2 presents summary statistics as well as separate statistics for between- and within-county variation for the variables. It is important to note that there is relatively little within-county variation for the agricultural variables.

4. Reduced Form Model

We estimate a discrete choice model by regressing the probability of investment in an ethanol plant on the covariates. First we estimate a fixed-effects logit model:

$$(1) \quad I_{it} = \alpha_i + a_{it}'\beta + X_{it}'\delta + Time_t'\gamma + \varepsilon_{it},$$

It is logical to begin with the fixed effects model since unobservable county characteristics might explain their ability and desire to invest in a plant. In this model a_{it} is the number of ethanol plants open at the start of a period – both in the county (*other plants*), and in the region (*spatial lag other plant*), -- and thus the direct competition for an entering plant. Since a_{it} represents the number of ethanol plants open at the start of the period in which the investment decision is made, it is not endogenous. X_{it} is a vector of exogenous covariates, and $Time_t$ is either a trend variable or a vector of year dummies, and ε_{it} is the error term. We also estimate a similar pooled logit model:

$$(2)\; I_{it} = \alpha + a_{it}'\beta + X_{it}'\delta + Time_t'\gamma + \varepsilon_{it}$$

which is identical to (1) except that it omits the county fixed effects.

In both (1) and (2), the vector X_{it} contains the exogenous covariates *biodiesel plant* (a count of plants in the county), *soy price, corn price, ethanol price, oil price, natural gas price, corn intensity* (fraction of county planted to corn), *spatial-lag corn intensity* (fraction of region, not including county, planted to corn), *soy intensity, cow density* (number of cows per county-acre), *gas terminal, ethanol terminal* (presence of terminal in county – for the pooled model only), *subsidy* (whether state has producer tax credit program), and *mtbe ban* (indicator of state MTBE ban in effect).

Each of these models has its advantages. The major advantage of the fixed-effects model (1) is the addition of the county specific effect. This allows us to control for unobservable county traits, such as openness to business, that remain fixed over time. This is particularly important since the resolution of our data is not ideal and some variables are not observed at the county level. The downside to the fixed-effects model is that it exploits within-group variation, so we can not identify time-invariant regressors, whose impacts are absorbed by the fixed effect. In the pooled logit model (2), pooling the errors over counties allows us to correct for some of the potential heteroskedasticity and autocorrelation present in panel models.

Therefore, in the fixed-effects regression (1), we only use observations for counties that experienced a new investment during our time frame. Conversely, in the pooled regression (2), we use all observations and are able to identify the time-invariant regressors since this estimation technique exploits variation across counties as well as over time. We also rejected a random effects model, and a Hausman test indicated that county unobservables are likely correlated with the regressors, indicating fixed effects is the appropriate specification.[2]

The results of the estimation of (1) and (2) are presented in tables 3 and 4, respectively. As a robustness check, in both cases we also include specifications where we run the regressions with a time-lagged corn intensity variable instead of contemporaneous corn intensity in order to control for potential endogeneity. We find no significant difference whether corn intensity is lagged or not. We do not anticipate endogeneity problems with the other variables since they are observed on a more aggregate level, and thus would not be expected to respond to the addition of one ethanol plant at the county level.

We also estimate (1) and (2) using a subsample of our data, eliminating data from states with fewer plants (Ohio, Missouri, Indiana, and Wisconsin). Since there are only 96 new plants in our sample, we have a potential problem with excess zeros. This sub-sample thus provides a robustness check whose results are qualitatively similar to the full sample.

[2] We obtained Chi2 Hausman statistics of 61.5 and 61.0 for testing between the base random effects and fixed effects models with year and year dummies. These allow us to reject the null hypothesis at a 1% significance level.

In the pooled model (2), we find that the effect of *other plant* was significant and negative. In some specifications we also found that the *spatial-lag other plant* had a significant and negative effect on entry, albeit smaller than the local effect. Taken together, this indicates there is a negative competitive effect between firms, and that the effect decreases with distance.

The data also indicates that availability of corn, via *corn intensity*, has a positive and significant impact on investment in ethanol plants. This effect is slightly smaller with the time-lagged variable. The regional variable is not significant, indicating ability to source locally is important. The *cow density* also has a significant and positive effect on investment indicating that the potential market for ethanol co-products is important (versus cows competing with plants for corn).

The *subsidy* variable is marginally significant and positive. We define this as a binary variable because the actual program of producer tax breaks and support varies by state, thus *subsidy* is just an indicator of the presence of some support for some plants in that state. More specifically, some states only support small or farmer owned plant, others only large plants; these differences decrease the precision of our estimate. We also note that more targeted analysis of the effects of these policies is the subject of on-going work.

Finally, we must discuss the effects of prices in the context of time. For regressions 1 and 2 in Table 3, we find that corn price has a negative and significant effect on investment, which makes sense since corn is the largest variable cost of operating a plant.

It is important to note that these two specifications incorporate time as year dummies. No price variables are significant in the specification with a time trend. This may be because the trend absorbs all the effects – oil and ethanol price do not vary spatially and are correlated, thus it is difficult to separate the effect of one from the other (though removing them from the regression has no effect on the other estimates). We still think it is worthwhile to include the trend variable because there have been technological changes in plants that affects the operation and construction costs.

The only similarity between the pooled model (Table 4) and fixed effects model (Table 3) is that *other plant* was highly significant and had a negative effect in all specifications. The only other significant variables in the fixed effects model were the time indicators. That said, these specifications have very high pseudo-R^2 values, indicating a good fit. The fixed effects themselves are highly significant and absorb the effect of the other observed and unobserved covariates. More specifically, many of our variables do not vary at the county level, so when we add the county-specific effects, they absorb the effect of those unobserved county-level covariates. This also indicates that county-level (location specific) intangibles are very important in determining where ethanol plants locate.

Finally, we also estimate (1) and (2) specifying the strategic variable, other plants, as capacity instead of count. We hypothesize that larger or smaller competing capacities could have different effects. The results were qualitatively the same for the fixed effects regression (Table 3, equations 5 and 6). However, we find differences between the two specifications of the plant competition in the pooled model (Table 4, equations 5 and 6). Specifically, other plant capacity was barely significant or insignificant whereas other plant was highly significant. The negative sign on the estimates indicates that as competing plants get larger, the deterrent effect also increases, which we would expect for competition in the input or output markets. Detangling these effects is part of ongoing analysis.

We estimate a third model as a robustness check of the investment decision in (1). In this case, we specify the plant investment variable as IC_{it}, such that

$$(3)\quad IC_{it} = \alpha + a_{it}'\beta + X_{it}'\delta + Time_t'\gamma + \varepsilon_{it},$$

where the continuous dependent variable IC_{it} is the capacity of the new plant. We hypothesize that the magnitude of certain effects might vary with respect to determining presence of the investment versus the size of that investment. We estimate this model using a tobit model with pooled errors.

The results for the tobit estimation (Table 5) are qualitative similar to the pooled logit results. Significantly, the effects of the *other plant* variable and for *corn intensity* are highly significant in the tobit model. This makes sense as far as competition with other plants is expected to be stronger if the larger plants use more corn and produce more ethanol. There is an important difference in this model, as we look at a size response to competition, and find that the presence of another plant leads to larger entering plants (Table 5, equations 1 and 2). We also find, however, that size of local existing plants does not affect size of entering plant, but the spatially lagged competing plants may have a negative effect (Table 5, equations 3 and 4). We follow this line in further work to disentangle the source of competition between plants.

5. Structural Model

We follow up the reduced form model of investment in an ethanol plant with a structural model for several reasons. First, it is interesting to estimate the effect of the state variable the expected payoff from investing in an ethanol plant. In the reduced form model, we estimated the effect of these variables on the probability of investment. Similarly, and importantly, the structural model makes it possible to estimate the strategic interaction between different regions that produce corn and could invest in ethanol plants.

Secondly, since location matters in determining the probability of an ethanol plant being built, as well as the ethanol plant's profitability, there may be spatial correlation in the unobserved variables. This problem is handled in the structural model because profits can be interpreted as expected profits, conditional on the observables (Pakes, Ostrovsky and Berry, 2005; Lin, 2010).

Thirdly, the structural model explicitly models the dynamic investment decision, including the continuation value to waiting. In contrast, the reduced-form model only estimates the per-period probability of investment.

Since dynamic games are static games where the decisions in any period are linked by the evolution of the state variables, it makes sense to begin this model by thinking about a static game. Bajari et al. (2006) describe a semi-parametric method for estimating static games of incomplete information. This model begins with a multinomial logit discrete choice problem (i.e. a random utility model), but also allows utility to depend on the actions of other agents.

Many of the first applications of discrete choice models to dynamic decisions only describe the actions of a single agent. Rust (1987) described and estimated and infinite horizon discrete choice problem using a nested fixed point maximum likelihood estimator. This model allowed him to solve for structural parameters without a closed form solution. Rothwell and Rust (1997) built on this model for a finite horizon game that allowed for more than a binary action choice in the operation of nuclear power plants.

Recent papers have developed techniques for estimating dynamic games between multiple agents. Pakes, Ostrovsky, and Berry (2005) illustrated a dynamic entry/exit game where the structural parameters could be estimated semi-parametrically. Bajari, Benkard, and Levin (2006) added to this literature by describing the estimation of the parameters in a dynamic game with continuous control variables.

This model follows Lin's (2010) structural model of a discrete dynamic investment timing game in the offshore petroleum industry. Lin builds on previous literature on discrete games of entry and exit by examining sequential investments with a finite-horizon. Her model allows the estimation of structural parameters and testing for strategic interactions using generalized method of moments. The model also accounts for spatial relationships among the firms making the investment decisions, an aspect crucial in our model as well.

In our structural model we thus estimate the theoretical model presented in section 3 using semi-parametric techniques. As in the theoretical model, the investment decision in each county *i* in year *t* depends on the state of the county $\Omega_{it} = (a_{it}, X_{it})$ through its effect on profits. The state variables a_{it} and X_{it} evolve according to a first-order Markov process and summarize the direct effect of the past on the current environment.

The endogenous state variable a_{it} captures the strategic components of the investment decision. In this model, a_{it} is a count of the number of ethanol plants that exist in a given county *i* at time *t*, where $a_{it} \in \{0,1,2\}$. The exogenous state variables in X_{it} are a subset of the variables used in the reduced form estimation. X_{it} includes *corn intensity*, *regional corn intensity, cow density, corn price, ethanol price, oil price, natural gas price, biodiesel plant, subsidy, and mtbe ban*. Due to the computational nature of the problem it is not possible to include all the variables in one specification, so we include those that we believe are important based on our reduced-form results.

In addition, the estimation technique requires that we discretize the state variables (due to dimensionality, we place each variable into two or three bins); we present summary statistics in Table 6.

The maximum number of *other plants* in a county is 1, so this variable is already binary in nature. For the *regional other plants* variable, there are up to 5 plants; we break this variable into three categories, for 0, 1, and more than 1 other plant in the region. There is a maximum of two biodiesel plants per county, but due to dimensionality, we describe the presence of *biodiesel plants* as either 0, or 1 if they are any plants. Likewise, the variables describing *mtbe ban,* and *subsidy* are already binary in nature.

We discretize the continuous variables into 3 bins centered on their means, where we define the bin widths to maximize both spatial and inter-temporal variation. As such, the middle bin for the *corn intensity*, *regional corn intensity,* and *cow density* variables are two standard deviations around the mean. In addition, we used an additional discretization of *regional corn intensity* with only 2 bins in order to decrease dimensionality.

For *corn price* and *natural gas price,* the middle bin is two standard deviations centered on the mean. Since there was more variation in *ethanol price*, we tried two discretization strategies. In the first, the middle bin is one-half standard deviation around the mean, and in the second the middle bin is two standard deviations around the mean. Finally, the oil price bins are not centered on the mean; instead the highest 2 years are the top bins, while prices below 35 dollars per barrel are the lowest bin.

In addition to the observable state variables, the expected profit from investing in an ethanol plant in each county *i* in year *t* depends on a private shock, ε_{it}. This error term captures information known in county *i*, but unknown to the other counties or the econometrician. It can be interpreted as a shock to the cost of building an ethanol plant. We assume the error term is distributed exponentially with parameter σ.

The profit from investing in an ethanol plant in county *i* in year *t*, $\Pi\left(a_{it}, X_{it}, \varepsilon_{it}; \theta\right)$, has the additively separable representation

$$\Pi(a_{it}, X_{it}, \varepsilon_{it}; \theta) = \Pi_o(a_{it}, X_{it}; \theta) + \varepsilon_{it}(a_{it}).$$

Additionally, we specify a linear utility function, such that

$$\Pi_o(a_{it}, X_{it}; \theta) = a_{it}'\gamma_a + X_{it}'\gamma_X,$$

where $\theta = (\gamma_a, \gamma_X)$ denotes the parameters to be estimated. The coefficient on a_{it} represents the effect of investment decisions in neighboring counties on the profit from investing in an ethanol plant. The coefficients on the exogenous state variables, X_{it}, show the effect of the state of the world on the profitability of an ethanol plant.

The value function for a county i in a given period t can be written as

$$V\left(a_{it}, X_{it}, \varepsilon_{it}; \theta\right) = \max\left\{\Pi\left(a_{it}, X_{it}, \varepsilon_{it}; \theta\right), \beta V^c\left(a_{it}, X_{it}; \theta\right)\right\}$$

The payoff in the county will depend on whether or not they decide to wait and not build an ethanol plant in year *t*, indicated by $I_{it} = 0$, or build an ethanol plant, $I_{it} = 1$. If they choose to build an ethanol plant in year t , they will receive the payoff $\prod(a_{it}, X_{it}, \varepsilon_{it}; \theta)$. If they choose not to invest at time t, then they receive the discounted continuation value, where β is a discount rate, and $V^c(\bullet)$ is the continuation value to waiting. Whether or not there is a new investment depends on which of these options yield the highest payoff in that particular period.

The continuation value $V^c(\bullet)$ is the expected value of the next period's value function, conditional on not building an ethanol plant in the current period, and is given by

$$V^c\left(a_{it}, X_{it}; \theta\right) = E\left[V\left(a_{it+1}, X_{it+1}, \varepsilon_{it+1}; \theta\right) | a_{it}, X_{it}, I_{it} = 0\right].$$

In this case, we assume the error term ε_{it} is distributed exponentially, so the continuation value reduces to:

$$V^c(a_{it}, X_{it}; \theta) = E[\beta V\left(a_{it+1}, X_{it+1}; \theta\right) + \sigma g\left(a_{it+1}, X_{it+1}; \theta\right) | a_{it}, X_{it}, I_{it} = 0]$$

as shown by Lin (2010). Here, σ is the parameter of the exponential error term, which is estimated along with the γ parameters.

The probability of investment in an ethanol plant in county i in year t, conditional on the state of the world, is given by

$$g(a_{it}, X_{it}; \theta) = \exp\left(-\frac{\beta V^C(a_{it}, X_{it}; \theta) - \Pi_o(a_{it}, X_{it}; \theta)}{\sigma}\right)$$

This policy function can also be interpreted as the perception from the point of view of one investor of the probability of another ethanol plant being built in the same county in a given state of the world.

We estimate the parameters using a two-step semi-parametric procedure. The first step estimates the continuation value non-parametrically, and the second step estimates the parameters $\theta = (\gamma_a, \gamma_x, \sigma)$ using generalized method of moments (GMM).

The first step in the estimation is to calculate the transition matrix M, where $M_{ij} = \Pr(\Omega_{it+1} = j \mid \Omega_{it} = i, I_{it} = 0)$. This describes the evolution of the state variables a_{it} and X_{it} over time, and can be calculated as an empirical average. Additionally, it is necessary to calculate the empirical probabilities of investment in an ethanol plant, $\overline{g}(a_{it}, X_{it}) = \Pr(I_{it} = 1 \mid \Omega_{it} = j)$ using empirical averages as well. $\overline{g}$ describes the probability of new investment in an ethanol plant for every possible state of the world.

The continuation value is specified as $\overline{V}^C = M\left(\beta \overline{V}_{t+1}^C + \sigma \overline{g}_{t+1}\right)$

where M is the transition matrix calculated previously, β is the discount rate, and $\overline{g}_{t+1}$ are the empirical probabilities also calculated previously. Since this is an infinite horizon problem, we first solve for a fixed point for the continuation value so that $\overline{V}^C$ and $\overline{V}_{t+1}^C$ are arbitrarily close. We then form the predicted probability of investment in an ethanol plant,

$$\widehat{g}(a_{it}, X_{it}; \theta) = -\frac{\beta \widehat{V}^C - \gamma_a a_{it} - X_{it}^{'} \gamma_X}{\sigma}$$

using this estimate for $\widehat{V}^C(a_{it}, X_{it}; \theta)$.

The second step of the estimation procedure moves on to the actual estimation of the parameters using GMM. The moment function is defined as:

$$\psi = \left(\left(\hat{g}(a_{it}, X_{it}; \theta) \right) - \overline{g}(a_{it}, X_{it}) \right) n(a_{it}, X_{it})$$

the difference between the predicted and empirical probabilities of investment in ethanol plants.. The data is incorporated into the model through the vector n, where n counts the number of times each state, $\Omega_{it} \mid I_{it-1} = 0$, occurs. Thus, ψ is a vector where each row represents difference in the predicted and empirical probabilities of investment in ethanol plants for each of the possible states of the world Ω_{it}, and is weighted by the number of times that state occurs in the data. The population moment condition is that in expectation, ψ equals zero.

Since $\theta = (\gamma_a, \gamma_X, \sigma)$, there must be at least one moment conditions for each parameter to be estimated. The GMM minimization problem to solve for the parameters is:

$$\min_{\theta} \left(\frac{1}{n\_obs} \sum \psi \right)^{'} W_n^{-1} \left(\frac{1}{n\_obs} \sum \psi \right),$$

where *n_obs* is the number of county-year observations. We create additional moments by interacting the moments ψ with the state variables Ω_{it}. Since the system is exactly identified, we substituted an identity matrix for the weight matrix W_n^{-1}. The robust standard errors for the estimates were obtained by bootstrapping with 100 replications, sampling counties with replacement.

We report the results of the structural estimation in Table 7. Note that due to dimensionality we were not able to include all covariates in one specification, thus we report several specifications that contain different combinations of covariates.

The presence of *other plants* in the county had a consistently large and negative effect on expected profits from entry. This indicates that strategic considerations are indeed important in the corn-ethanol market. In this model, the *regional other plants* did not significantly affect profit, indicating that plants compete for inputs (or to sell outputs) locally, hence plants that are farther away do not affect expected profits.

The other covariate that has a consistently large significant effect on expected profits is the *mtbe ban*. Since ethanol is difficult and costly to transport, one might expect the ban, and thus local and regional demand for ethanol as an additive, would positively affect expected profits, both via lower transport cost and higher local price. Conversely, the producer support was not significant. Since the tax credit themselves are only applicable in some situations, they may not affect the expected profits of a typical plant.

The states of the input and output markets were important, but the significance of these effects varied across specifications due to collinearity. In general, *ethanol price* had a significant and positive effect, and *oil price* a larger negative effect on expected profits.

This confirms that oil and ethanol are mainly complements, thus demand for ethanol is closely tied to oil demand and usage. We should also note that the resolution of these two prices was not ideal, as we only we able to exploit intertemporal variation.

Likewise, the *corn price* and *ethanol price* were generally significant and had a negative effect on expected profits, as would be expected of inputs. As with the fuel prices, the precision of these estimated could potentially be improved with better resolution of the data, which is currently only available at the state level.

Overall, the corn intensity variables were positive, and either *regional corn intensity* or the local *corn intensity* were significant. Taken together, this indicates that availability of inputs is important in determining expected profits from investment in an ethanol plant. We may not have been able to separate the local and regional effect because certain regions have higher production, and variation across those regions may have been smoothed out by binning the variables.

6. Conclusion

The results of both the reduced form and structural estimation indicate that there is an important strategic component to investment in ethanol plants. There is a negative competitive effect between ethanol plants, though within this framework we can not further differentiate between competition in the inputs (corn) versus output markets.

We also find that intensity of corn production is important in determining local investment in both models. Corn is bulky and transportation is not cheap, so it is beneficial for plants to locate where they have good access to feedstock.

We find mixed results of the effects of input and output prices across the different models and specifications. This inconsistency is potentially due to the data resolution, which, at state and national level, is not ideal. That said, we were still able to find some effect of prices indicating that 1) they do matter, and 2) even with less than ideal data source, our model is strong enough to tease out some of these effects.

We also find mixed results of the policy variables. In the reduced-form specification, we find that the state producer tax breaks are important, and the MTBE ban is significant in the structural model. The differences here may be due to the models themselves. The MTBE ban affects the market via increased ethanol demand and higher expected prices; thus in the structural model it leads to increased expected profits from investment. The producer support/tax break only applies to some plants, and only to some of the production, so we may not have identified this effect in the structural model, where the profit function is the same for each county. Furthur analysis of the effects of policies on investment in ethanol plants is subject of ongoing work.

References

Bajari, P., C. L. Benkard, J. Levin (2007). "Estimating Dynamic Models of Imperfect Competition". *Econometrica*, 75(5): 1331-70.

Deparment of Energy (2008). Energy Time Lines: Ethanol. Revised June, 2008. Accessed January 2009.

Dhuyvetter, K. C., T. L. Kastens, M. Boland (2005). The US Ethanol Industry: Where Will It Be Located in the Future? Agricultural Marketing Resource Center and Agricultural Issues Center, University of California.

Federal Trade Commission (2008). Report on Ethanol Market Concentration.

Goetz, S. (1997). "State- and County-Level Determinants of Food Manufacturing Establishment Growth: 1987-93". American Journal of Agricultural Economics, 79: 838-850.

Lambert, D. M., M. Wilcox, A. English, L. Stewart (2008). "Ethanol Plant Location Determinants and County Comparative Advantage". Journal of Agricultural and Applied *Economics,* 40:117-135.

Lin, C.-Y. C. (2010). Do firms interact strategically?: A structural model of the multi-stage investment timing game in offshore petroleum production. In submission. University of California-Davis.

Low, S. A., and A. M. Isserman (2009). “Ethanol and the Local Economy: Industry Trends, Location Factors, Economic Impacts, and Risks”. *Economic Development Quarterly,* 23: 71-88.

McNew, K., and D. Griffith (2005). “Measuring the Impact of Ethanol Plants on Local Grain Prices”. *Review of Agricultural Economic,s* 27(2): 16-180.

Pakes, A., M. Ostrovsky, S. Berry (2007). “Simple estimators for the parameters of discrete dynamic games (with entry and exit examples)”. *RAND Journal of Economics,* 38(2): 373.

Rothwell, G., J. Rust (1997). “On the optimal lifetime of nuclear power plants”. *Journal of Business and Economic Statistics*, 15(2):195-208.

Rust, J. (1987). “Optimal replacement of GMC bus engines: an empirical model of Harold Zurcher”. *Econometrica*, 55(5): 999-1033.

Sarmiento, C., and W. W. Wilson (2007). "Spatial Competition And Ethanol Plant Location Decisions," 2008 Annual Meeting, July 27-29, 2008, Orlando, Florida 6175, American Agricultural Economics Association.

Urbanchuk, J. M. (2006). “Economic Impacts on the Farm Community of Cooperative Ownership of Ethanol Production”. National Corn Growers Association Report.

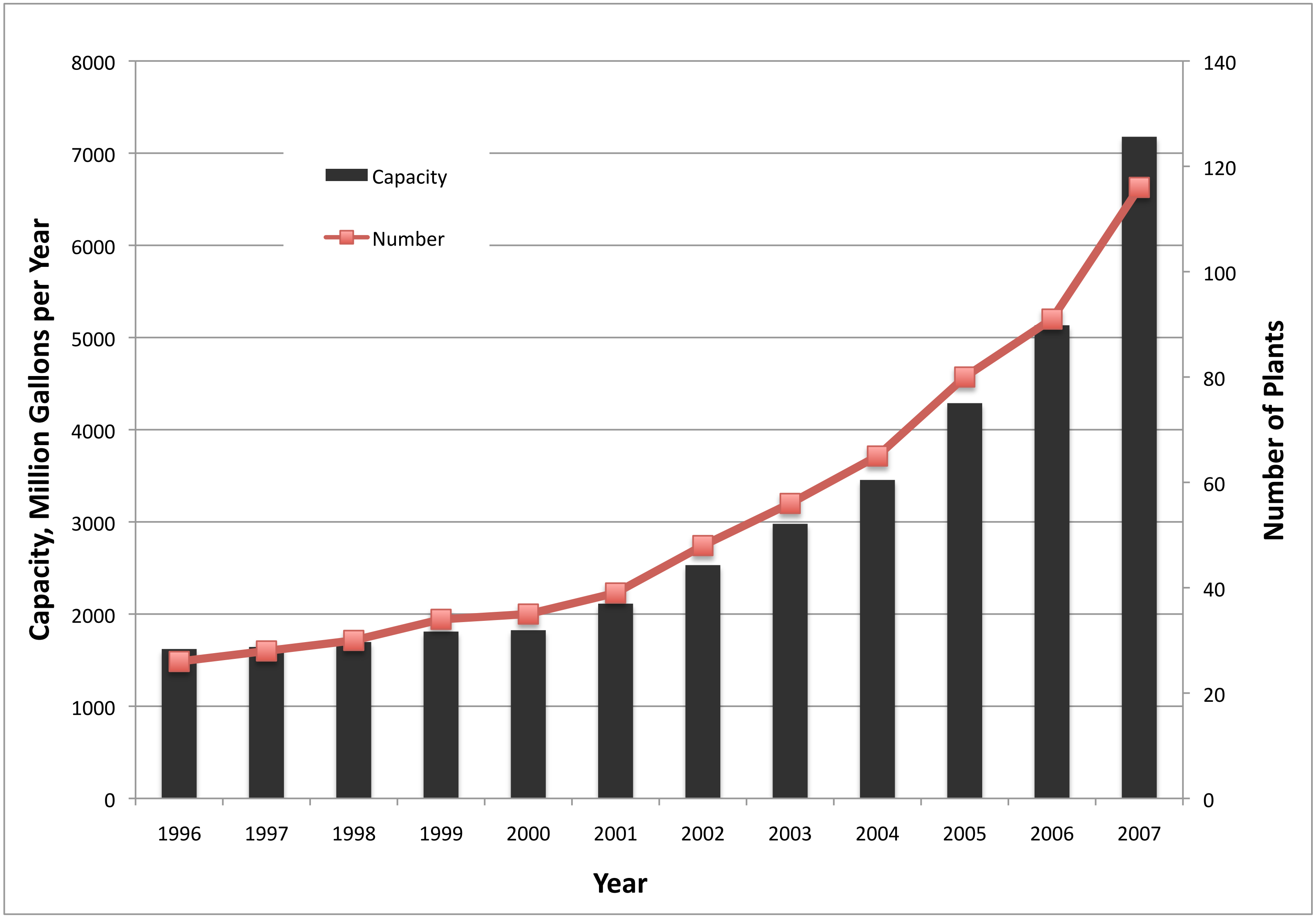

Figure 1:Plants and Capacity of the US-Midwest Ethanol Industry during the modern boom

Table 1: Summary statistics for full sample and sub-samples of counties with and without plants

VARIABLE	FULL SAMPLE MEAN	FULL SAMPLE STD. DEV.	COUNTIES WITHOUT OPEN PLANT MEAN	COUNTIES WITHOUT OPEN PLANT STD. DEV.	COUNTIES WITH OPEN PLANT MEAN	COUNTIES WITH OPEN PLANT STD. DEV.	COUNTIES WITH NEW INVESTMENT MEAN	COUNTIES WITH NEW INVESTMENT STD. DEV.
New Plant*	0.009	0.095	0	0	0.155	0.362	1.000	0.000
No. Plants at Start of Year	0.053	0.237	0	0	0.899	0.439	0.042	0.202
Capacity at Start of Year	2.930	18.639	0	0	49.786	59.795	1.905	12.411
Regional No. Plants at Start of Year	0.341	0.724	0.318	0.698	0.716	0.984	0.695	1.011
Regional Capacity at Start of Year	18.987	53.873	17.686	52.781	39.791	65.662	34.621	58.562
No. Biodiesel Plants	0.012	0.110	0.010	0.103	0.033	0.187	0.042	0.249
Gas Terminal	0.126	0.331	0.119	0.324	0.230	0.421	0.168	0.376
Ethanol Terminal	0.071	0.258	0.063	0.243	0.202	0.402	0.137	0.346
Regional Gas Terminal	0.787	0.797	0.777	0.790	0.941	0.889	0.853	0.799
Regional Ethanol Terminal	0.460	0.656	0.443	0.647	0.723	0.746	0.663	0.724
Soy Price	7.007	1.662	6.996	1.651	7.199	1.816	7.522	1.913
Corn Price	2.719	0.586	2.715	0.579	2.774	0.686	2.886	0.763
Ethanol Price	1.721	0.396	1.712	0.392	1.877	0.439	1.914	0.426
Oil Price	36.194	15.797	35.712	15.563	43.899	17.437	47.229	17.461
Natural Gas Price	6.288	1.900	6.251	1.903	6.892	1.740	7.140	1.677
Corn Intensity (Production)	27.556	23.988	26.154	23.156	49.544	26.028	49.592	26.139
Corn Intensity (Acreage)	0.200	0.148	0.192	0.145	0.330	0.138	0.325	0.135
Regional Corn Intensity (Acreage)	0.195	0.134	0.277	0.123	0.186	0.132	0.297	0.129
Regional Corn Intensity (Production)	26.904	21.623	38.716	20.817	25.536	21.296	45.081	24.595
Soy Intensity (Production)	7.847	6.061	7.642	6.042	10.884	5.507	10.833	6.049
Soy Intensity (Acreage)	0.189	0.132	0.186	0.132	0.245	0.112	0.239	0.124
Cow Density	0.985	0.665	0.972	0.663	1.183	0.653	1.208	0.698
Producer Support*	0.336	0.472	0.332	0.471	0.403	0.491	0.442	0.499
MTBE Ban*	0.438	0.496	0.420	0.494	0.739	0.440	0.800	0.402
Observations	10416		9803		613		95	

* Binary variables

Table 2: Summary statistics and between and within variation for selected variables.

Variable		Mean	Std. Dev
New Plant	overall	0.011	0.104
	between		0.028
	within		0.100
Other Plant	overall	0.056	0.243
	between		0.205
	within		0.131
Other Capacity	overall	3.149	18.942
	between		17.343
	within		7.637
Regional Other Plant	overall	0.361	0.766
	between		0.623
	within		0.446
Regional Other Capacity	overall	20.482	56.286
	between		49.630
	within		26.599
Biodiesel Plant	overall	0.014	0.124
	between		0.082
	within		0.093
Soy Price	overall	7.006	1.659
	between		0.153
	within		1.652
Corn Price	overall	2.718	0.585
	between		0.110
	within		0.574
Natural Gas Price	overall	6.430	1.908
	between		0.497
	within		1.842
Corn Intensity	overall	0.199	0.149
	between		0.149
	within		0.020
Regional Corn Intensity	overall	0.186	0.137
	between		0.136
	within		0.016
Soy Intensity	overall	0.188	0.132
	between		0.133
	within		0.022
Cow Density	overall	0.976	0.667
	between		0.664
	within		0.061
Subsidy	overall	0.350	0.477
	between		0.263
	within		0.398
MTBE Ban	overall	0.478	0.500
	between		0.159
	within		0.474

Table 3: Results from Fixed Effects Logit Estimation of Investment in New Ethanol Plants in the Midwestern United States

	Whole Sample, 10 States						Small Sample, 6 States#	
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(1)	(2)
Other Plant	-15.54***	-15.42***	-14.60***	-14.45***			-14.15***	-13.48***
	(2.070)	(1.965)	(2.001)	(1.902)			(2.069)	(1.842)
Other Plant Capacity					-0.233***	-0.238***		
					(0.0326)	(0.0334)		
Spatial Lag Other Plant	0.229	0.254	0.160	0.161			0.178	0.272
	(0.651)	(0.611)	(0.604)	(0.562)			(0.648)	(0.592)
Spatial Lag Other Plant Capacity					-0.00212	-0.00428		
					(0.00688)	(0.00649)		
Biodiesel Plant	0.979	1.041	0.475	0.649	1.513	1.655	-11.14	-10.54
	(2.771)	(2.613)	(1.910)	(1.835)	(2.170)	(2.015)	(1470)	(1017)
Soy Price	-1.584	-0.383	-1.092	-0.382	-1.076	-0.103	-0.202	-0.344
	(1.427)	(0.354)	(1.453)	(0.374)	(1.002)	(0.266)	(1.559)	(0.374)
Corn Price	0.0889	1.958	-1.610	1.674	-4.583	0.581	6.488	1.488
	(4.374)	(1.329)	(4.536)	(1.347)	(3.623)	(0.968)	(6.196)	(1.402)
Natural Gas Price	-0.158	0.0893	0.125	0.0438	0.0507	-0.0379	-0.201	-0.0586
	(1.246)	(0.365)	(1.236)	(0.388)	(0.834)	(0.288)	(1.328)	(0.394)
Ethanol Price		-1.480		-1.704		-0.947		-1.354
		(1.485)		(1.621)		(1.140)		(1.491)
Oil Price		0.0855		0.0903		0.0756		0.100
		(0.0511)		(0.0525)		(0.0389)		(0.0564)
Corn Intensity	18.96	17.50			2.919	2.971	11.96	11.43
	(20.55)	(20.53)			(15.51)	(15.58)	(21.28)	(20.20)
Time Lag Corn Intensity			2.294	6.489				
			(15.42)	(13.68)				
Spatial Lag Corn Intensity	-13.88	-12.88	3.475	-0.0148	-2.504	-9.563	-31.96	-21.93
	(29.06)	(25.86)	(26.00)	(22.05)	(22.32)	(20.72)	(31.27)	(26.37)
Soy Intensity	2.270	-1.272	-3.105	-6.459	-4.272	-9.399	-6.308	-6.693
	(13.76)	(13.49)	(14.45)	(13.90)	(9.798)	(9.850)	(14.62)	(13.77)
Cow Density	2.063	1.922	1.003	1.192	0.805	0.205	0.712	1.581
	(4.333)	(4.149)	(4.340)	(4.296)	(3.086)	(2.952)	(4.423)	(4.030)
Subsidy	0.798	0.922	0.853	0.949	1.051	1.031*	1.611	1.609*
	(0.741)	(0.693)	(0.827)	(0.745)	(0.547)	(0.513)	(0.887)	(0.765)
MTBE Ban	-0.604	-0.801	-0.527	-0.976	-0.268	-0.957	0.164	-0.773
	(1.139)	(0.884)	(1.140)	(0.912)	(0.846)	(0.629)	(1.357)	(0.962)
Year		1.074***		1.175***		0.697***		1.023***
		(0.261)		(0.287)		(0.179)		(0.272)
Year Dummies	YES		YES		YES		YES	
(Joint Significance Chi2)	37.6***		29.6**		39.8***		30.6**	
Observations	1098	1098	964	964	1098	1098	882	882
Number of Counties	92	92	88	88	92	92	74	74
Chi2 for Regression	354.6***	348.1***	325.6***	319.6***	270.9***	261.1***	278.2***	268.5***
pseudo-R2	0.766	0.752	0.761	0.747	0.585	0.564	0.745	0.719

Robust standard errors in parentheses

*** p<0.001, ** p<0.01, * p<0.05

Full sample less Ohio, Indiana, Missouri, and Wisconsin

Table 4: Results from Pooled Logit Estimation of Investment in New Ethanol Plants in the Midwestern United States

	Whole Sample, 10 States						Small Sample, 6 States#	
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(1)	(2)
Other Plant	-1.744**	-1.415**	-1.756**	-1.377**			-1.665**	-1.318*
	(0.558)	(0.541)	(0.551)	(0.529)			(0.555)	(0.531)
Other Plant Capacity					-0.0353*	-0.0279		
					(0.0175)	(0.0157)		
Spatial Lag Other Plant	-0.346**	-0.154	-0.402**	-0.188			-0.343*	-0.157
	(0.129)	(0.116)	(0.140)	(0.127)			(0.139)	(0.123)
Spatial Lag Other Plant Capacity					-0.00609*	-0.00407*		
					(0.00251)	(0.00191)		
Biodiesel Plant	0.686	0.531	0.555	0.401	0.561	0.407	-0.513	-0.596
	(0.616)	(0.588)	(0.655)	(0.641)	(0.645)	(0.630)	(0.970)	(0.937)
Soy Price	-1.090*	0.154	-0.934	0.102	-1.017	0.146	-1.509*	0.185
	(0.548)	(0.207)	(0.538)	(0.231)	(0.541)	(0.206)	(0.696)	(0.227)
Corn Price	-5.235***	-0.642	-5.956***	-0.368	-4.690***	-0.590	-4.253**	-0.710
	(1.401)	(0.685)	(1.474)	(0.743)	(1.351)	(0.679)	(1.382)	(0.780)
Natural Gas Price	0.0262	-0.177	0.0739	-0.142	0.0182	-0.166	0.743	-0.150
	(0.233)	(0.167)	(0.234)	(0.172)	(0.236)	(0.166)	(0.519)	(0.213)
Ethanol Price		-0.505		-1.158		-0.542		-0.412
		(0.811)		(0.936)		(0.817)		(0.895)
Oil Price		0.0567**		0.0698**		0.0580**		0.0559*
		(0.0219)		(0.0233)		(0.0217)		(0.0243)
Corn Intensity	8.613***	7.282***			8.704***	7.523***	8.052***	6.785***
	(1.748)	(1.570)			(1.768)	(1.594)	(1.967)	(1.839)
Time Lag Corn Intensity			7.859***	6.488***				
			(1.771)	(1.665)				
Spatial Lag Corn Intensity	-0.398	-1.316	0.232	-0.862	-0.679	-1.348	-0.563	-2.109
	(2.290)	(2.075)	(2.335)	(2.178)	(2.367)	(2.158)	(2.670)	(2.497)
Soy Intensity	-1.150	-0.281	-0.670	0.234	-1.024	-0.301	-0.653	0.691
	(1.424)	(1.292)	(1.521)	(1.367)	(1.470)	(1.319)	(1.751)	(1.589)
Cow Density	0.354*	0.425**	0.384*	0.429**	0.363*	0.422**	0.191	0.322*
	(0.163)	(0.132)	(0.162)	(0.135)	(0.158)	(0.129)	(0.195)	(0.150)
Subsidy	0.624*	0.648*	0.597*	0.567*	0.579*	0.606*	0.485	0.635*
	(0.246)	(0.261)	(0.259)	(0.278)	(0.246)	(0.259)	(0.293)	(0.317)
MTBE Ban	0.139	0.875*	0.0550	0.924*	0.109	0.848*	0.233	0.900*
	(0.461)	(0.415)	(0.468)	(0.428)	(0.462)	(0.412)	(0.788)	(0.445)
Gas Terminal			0.255	0.0613	0.202	0.0526		
			(0.751)	(0.760)	(0.769)	(0.762)		
Ethanol Terminal			0.383	0.574	0.346	0.528		
			(0.823)	(0.832)	(0.844)	(0.841)		
Year		-0.00344***		-0.00344***		-0.00354***		-0.00342***
		(0.000581)		(0.000605)		(0.000578)		(0.000686)
Year Dummies	YES		YES		YES		YES	
(Joint Significance Chi2)	144.8***		182.6***		163.5***		146.5***	
Observations	9745	9745	8960	8960	9745	9745	5970	5970
Chi2 for Regression	1446***	1618***	1431***	1552***	1465***	1641***	1152***	1274***

Robust standard errors in parentheses

*** p<0.001, ** p<0.01, * p<0.05

Full sample less Ohio, Indiana, Missouri, and Wisconsin

Figure 5: Tobit estimation of size of investment in new ethanol plants.

VARIABLES	(1)	(2)	(3)	(4)
Other Plant	33.43***	39.67***		
	(9.888)	(9.023)		
Other Plant Capacity			0.0875	0.106
			(0.143)	(0.135)
Spatial Lag Other Plant	-9.449	-4.488		
	(5.006)	(4.836)		
Spatial Lag Other Plant Capacity			-0.176*	-0.136
			(0.0861)	(0.0737)
Biodiesel Plant	26.40	20.57	26.55	22.55
	(21.41)	(21.51)	(21.49)	(21.53)
Soy Price	-41.71	-0.600	-39.51	-0.645
	(21.72)	(7.433)	(21.80)	(7.449)
Corn Price	-143.5**	0.971	-149.2**	-0.607
	(51.95)	(24.31)	(52.71)	(24.37)
Natural Gas Price	-1.398	-1.336	-2.538	-1.666
	(7.456)	(6.147)	(7.384)	(6.232)
Ethanol Price		-43.03		-40.47
		(28.83)		(28.23)
Oil Price		2.048*		2.128**
		(0.817)		(0.806)
Corn Intensity	261.7***	246.3***	303.4***	292.3***
	(73.89)	(71.12)	(71.91)	(70.21)
Spatial Lag Corn Intensity	26.43	-14.65	9.604	-19.89
	(79.72)	(78.04)	(82.19)	(80.73)
Soy Intensity	-42.79	-21.59	-54.91	-39.93
	(52.30)	(50.65)	(55.17)	(53.43)
Cow Density	13.15*	14.14*	13.77*	14.88**
	(6.188)	(5.579)	(6.373)	(5.731)
Subsidy	28.30**	31.38***	28.65**	32.19***
	(9.034)	(9.075)	(8.774)	(8.982)
MTBE Ban	5.230	36.02*	4.583	37.01*
	(15.66)	(14.41)	(15.76)	(14.43)
Gas Terminal	-3.716	-9.651	-4.320	-10.01
	(28.64)	(28.78)	(29.43)	(29.54)
Ethanol Terminal	27.02	35.42	30.84	39.19
	(30.99)	(31.06)	(32.20)	(32.40)
Year		-0.174***		-0.176***
		(0.0252)		(0.0251)
Year Dummies	YES		YES	
(Joint Significance Chi2)	33.0***		39.8***	
Observations	9745	9745	9745	9745

Robust standard errors in parentheses

*** p<0.001, ** p<0.01, * p<0.05

Table 6: Summary Statistics for Binned Variables

VARIABLE	FULL SAMPLE MEAN	FULL SAMPLE STD. DEV.	COUNTIES WITHOUT OPEN PLANT MEAN	COUNTIES WITHOUT OPEN PLANT STD. DEV.	COUNTIES WITH OPEN PLANT MEAN	COUNTIES WITH OPEN PLANT STD. DEV.	COUNTIES WITH NEW INVESTMENT MEAN	COUNTIES WITH NEW INVESTMENT STD. DEV.
New Plant	0.009	0.095	0	0	0.155	0.362	1.000	0.000
Other Plant	0.053	0.237	0	0	0.899	0.439	0.042	0.202
Regional Plants	0.308	0.599	0	1	0.633	0.765	0.611	0.816
Biodiesel Plant	0.012	0.110	0	0	0.033	0.187	0.042	0.249
Gas Terminal	0.126	0.331	0.119	0.324	0.230	0.421	0.168	0.376
Ethanol Terminal	0.071	0.258	0.063	0.243	0.202	0.402	0.137	0.346
Producer Support	0.336	0.472	0.332	0.471	0.403	0.491	0.442	0.499
MTBE Ban	0.438	0.496	0.420	0.494	0.739	0.440	0.800	0.402
Oil Price (2 bins)	0.333	0.471	0.320	0.467	0.545	0.498	0.632	0.485
Oil Price (3 bins)	1.167	0.553	1.152	0.548	1.398	0.576	1.505	0.543
Natural Gas Price	0.954	0.668	0.943	0.669	1.134	0.638	1.179	0.618
Corn Price	1.027	0.480	1.028	0.471	1.011	0.609	1.084	0.679
Ethanol Price A	0.917	0.759	0.897	0.754	1.228	0.780	1.337	0.780
Ethanol Price B	1.000	0.577	0.988	0.571	1.192	0.641	1.232	0.691
Corn Intensity	0.819	0.496	0.799	0.492	1.131	0.446	1.126	0.419
Regional Corn Intensity	0.985	0.708	0.952	0.701	1.515	0.598	1.495	0.581
Cow Density	0.926	0.647	0.913	0.651	1.132	0.531	1.179	0.525

Table 7: Results from Structural Estimation of Investment in Ethanol Plants

	A	B	C	D	E	F	G
Other Plant	-152.635*** (48.795)	-159.241*** (50.293)	-167.44*** (46.723)	-158.26** (75.995)	-155.316*** (57.317)	-161.764*** (56.819)	-154.484*** (36.628)
Regional Other Plant	40.575 (43.550)	101.096*** (40.171)	38.318 (38.158)	45.032 (42.042)	30.139 (51.557)	30.498 (56.827)	42.036 (43.022)
Corn Intensity	46.702** (22.606)	-2.429 (24.848)	10.435 (23.342)	9.5592 (18.128)	9.424 (27.419)	13.607 (24.909)	18.481 (23.293)
Regional Corn Intensity		46.199** (20.900)		56.799*** (22.381)	67.046** (30.581)	66.854** (28.210)	47.245*** (18.207)
Regional Corn Intensity 2	46.606 (36.619)		54.807 (36.826)				
Cow Density	21.72 (13.867)	9.481 (11.132)	28.045*** (10.382)	24.809** (12.384)			19.780 (13.302)
Corn Price	3.584 (7.397)	-21.401*** (8.280)	-31.859** (15.840)	-36.377*** (11.629)	3.834 (7.799)	-30.534*** (9.167)	-29.533*** (7.904)
Ethanol Price 1		3.161 (13.608)		-3.3283 (8.7197)			
Ethanol Price 2	17.139* (9.081)		0.830 (11.415)		28.962*** (7.353)	38.907*** (9.610)	25.238** (10.187)
Oil Price	-52.851*** (18.808)				-61.562*** (23.587)		
Natural Gas Price						-40.621*** (12.332)	-40.597*** (8.864)
MTBE Ban	159.115*** (53.047)	115.519*** (31.475)	121.65*** (35.151)	114.9* (63.55)	179.900*** (68.860)	155.957*** (59.099)	140.666*** (37.972)
Subsidy		15.536 (26.695)					
Biodiesel Plant			36.029 (74.902)	53.568 (60.866)			
Sigma	1999.0*** (0.346)	1998.6*** (0.378)	1999.6*** (0.118)	1998.9** (979.31)	1998.8*** (0.511)	1998.6*** (0.530)	1999.1*** (0.271)

Each Specification has 10,385 observations from 865 counties
Standard Errors in Parentheses
* p<10%, ** p<5%, *** p<1%