Corn Ethanol in the Midwestern United States: Local Competition, Entry, and Agglomeration*

Karen E. Thome C.-Y. Cynthia Lin Lawell

December 30, 2018

Abstract

Ethanol has attracted considerable attention both for its use as a gasoline substitute, and as a way to enhance profits in rural areas. In this paper, we analyze the entry of corn-ethanol plants in the Midwestern United States, where the majority of corn in the US is grown, during the second US ethanol boom. Results show that that the number of existing ethanol plants have a negative effect on the probability of ethanol plant entry in a given county. We also find evidence that the number of existing ethanol plants in the counties bordering a given county may have a positive effect on ethanol plant entry in that county. In particular, existing conglomerates and large ethanol producing firms in neighboring counties have a positive effect on ethanol plant entry, while existing singlet plants in neighboring counties do not. These results suggest that the net negative competitive effect among plants not only dissipates with distance, but also becomes net positive, indicating possible agglomeration benefits in the ethanol industry.

Keywords: ethanol, entry, local competition, agglomeration
JEL codes: Q16, L13

*Thome: USDA Economic Research Service; karen.thome@ers.usda.gov. Lin Lawell: Cornell University; clinlawell@cornell.edu. We thank Bruce Babcock, Colin Cameron, Ben Dawson, Timothy Fitzgerald, Scott Irwin, Marc Rysman, Rich Sexton, and Jim Stock for helpful discussions and comments. We also benefited from comments from conference participants at the Association of Environmental and Resource Economists Conference, the American Economic Association annual meeting, the International Industrial Organization Conference, the Berkeley Bioeconomy Conference, the Stanford Berkeley Industrial Organization Conference, and the Statistical Methodology in the Social Sciences conference; and seminar participants at Duke University and Rice University. We received funding from a UC-Davis Chevron Research Grant; a University of California at Davis Hellman Fellowship; and the National Center for Sustainable Transportation, which is supported by the U.S. Department of Transportation through the University Transportation Centers program. The views expressed are those of the authors and should not be attributed to the Economic Research Service or USDA. All errors are our own.

1 Introduction

Ethanol has attracted considerable attention both as an environmentally-friendly alternative to imported oil, and as a way to boost farm profits and improve rural livelihoods. Fuel ethanol can play different roles in the energy market, as an energy substitute for gasoline, or as an additive (oxygenate and/or octane booster) to gasoline. In the United States, a boom in the construction of corn-ethanol plants, known as the second US ethanol boom, began in the mid-1990s and hit full-stride by the early 2000s.[1]

This paper focuses on ethanol plant entry decisions in the Midwestern United States during the second US ethanol boom. In particular, we analyze how strategic interactions, government policy, and economic factors affect ethanol plant entry decisions over the period 1996-2008.

There are two sources of strategic interactions that add a strategic (or non-cooperative) dimension to potential entrants' entry decisions. The first is a competition effect: if there is more than one ethanol plant located in the same region, these plants may compete in the local feedstock input market or they may compete in the local fuel ethanol output market. The competition effect deters ethanol plants from entering in regions where there are other ethanol plants already present.

High transportation costs in both the feedstock and ethanol markets may be one reason for localized competition among neighboring plants. Empirical evidence has shown that industries with high transport costs are less geographically concentrated (Behrens, Brown and Bougna, 2018).

Feedstock is approximately 70% of the cost of producing corn-ethanol, and transportation costs for the bulky grains constitute a significant share (Whittington, 2006). As a consequence, the distance from a plant to the feedstock production area is extremely important. For example, Sarmiento, Wilson and Dahl (2012) find that competition in feedstock

[1]The first US ethanol boom stemmed from the desire for more energy self-sufficiency in the aftermath of the oil embargoes in 1973 and 1979, and led to the construction of 153 new plants by 1985 (DOE, 2008). For a more detailed discussion of the first and second US ethanol boom, see Lin Lawell (2017).

procurement can lead to a negative competition effect in localized corn markets, and that a shift in demand from a new plant could increase corn feedstock prices. Thus, owing to high transportation costs, neighboring ethanol plants may compete in the local feedstock input market.

Fuel ethanol transportation is more difficult, and thus is more expensive, than gasoline transportation because ethanol can easily absorb water during the transportation process. As a consequence, unlike gasoline, which can be transported via pipelines, fuel ethanol must be transported using specialized tank trucks. Rail is the primary form of transport used to ship ethanol from the Midwestern US to each coast. Rail transport has become increasingly congested given the growth in domestic crude oil production (EIA, 2015; Bushnell, Hughes and Smith, 2017) and the small number of firms that operate most national rail routes (Preonas, 2018). Neighboring ethanol plants may therefore compete over access to transportation for their ethanol output, leading to higher marketing costs for fuel ethanol. Thus, owing to high transportation costs, neighboring ethanol plants may also compete in the local fuel ethanol output market.

The second source of strategic interaction is an agglomeration effect (Goetz, 1997; Ellison and Glaeser, 1999; Zhu et al., 2011; Ahlfeldt et al., 2015; Gaubert, 2018; Michael Pflüger and Tabuchi, 2019; Verstraten, Verweij and Zwaneveld, forthcoming); if there are several ethanol plants located in the same region, the existing plants may have developed transportation and marketing infrastructure and/or an educated work force from which entering plants can benefit (Lambert et al., 2008).

Several government policies have coincided with the second US ethanol boom. First, the Clean Air Act Amendments of 1990 mandated the use of oxygenates, which include ethanol, in gasoline. The subsequent phase out and ban of the oxygenate methyl tertiary-butyl ether (MTBE) as a gasoline additive beginning in the late 1990s further increased the demand for ethanol. Second, the Renewable Fuel Standard (RFS) was created under the Energy Policy Act of 2005 with the goal of accelerating the use of fuels derived from renewable sources (EPA, 2013). The initial RFS (RFS1) mandated that a minimum of 4 billion gallons be used in 2006, rising to 7.5 billion gallons by 2012. Two years later, the Energy Independence and Security Act of 2007 greatly expanded the biofuel mandate volumes, creating the RFS2, which requires steadily increasing volumes of biofuel to be blended into the nation's fuel supply, reaching 37 billion gallons (bgal) a year by 2022. Third, many states have offered tax credits to ethanol producers (Cotti and Skidmore, 2010). These federal and state policies have coincided with increases in petroleum prices that made ethanol more competitive as an energy substitute for gasoline (Gallagher, 2009). Over this time period, the number of operational ethanol plants rose from 35 plants in 1991, to 50 plants in 1999, to 192 plants in September of 2010, for a total capacity of 13 billion gallons per year.

Our results show that that the number of existing ethanol plants has a negative effect on the probability of ethanol plant entry in a given county. We also find evidence that the number of existing ethanol plants in the counties bordering a given county may have a positive effect on entry in that county. In particular, existing conglomerates and large ethanol producing firms in neighboring counties have a positive effect on ethanol plant entry, while existing singlet plants in neighboring counties do not. These results suggest that the net negative competitive effect among plants not only dissipates with distance, but also becomes net positive, indicating possible agglomeration benefits in the ethanol industry.

The balance of our paper proceeds as follows. In Section 2, we review the relevant literature. We present our empirical model in Section 3. We describe our data in Section 4. We present our results in Section 5. Section 6 concludes.

2 Literature Review

2.1 Ethanol entry and location decisions

The first branch of literature on which we build is that on models of firm entry and location decisions. For excellent reviews of this literature, see Goetz (1997) and Bartik (1985). In empirical models of firm entry and location decisions, firm entry, particularly in manufacturing, is often modeled as a function of output market prices and access, input costs and access, and the policy environment. In some papers, such as Goetz (1997), location decisions involve a two-step process in which potential entrants first choose regions for broader consideration based on one set of criteria, and then narrow the choice within each region based on another set of criteria.

In the previous literature on ethanol plant location decisions, Sarmiento, Wilson and Dahl (2012) use a cross-sectional discrete choice model to analyze the agricultural characteristics and spatial dimensions that determine ethanol plant location, and find a large negative effect of a nearby plant on the probability of another plant locating nearby, and furthermore, that this effect decreases with distance. Similarly, Lambert et al. (2008) use a cross-sectional 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 find a negative impact on the location of plants that entered between 2000 and 2007. Haddad, Taylor and Owusu (2010) model state-by-state spatial determinants of plant location. Cotti and Skidmore (2010) estimate a model of investment in ethanol over time using aggregate state-level data on investments.

The results of these studies are not always qualitatively similar, however, because of the different empirical specifications and the different regions examined. For example, Sarmiento, Wilson and Dahl (2012) and Lambert et al. (2008) find that access to corn is an important location determinant. In contrast, Haddad, Taylor and Owusu (2010) do not find access to corn to be significant, though they note that following location theory (e.g. Goetz, 1997), firms might first choose a region with a lot of corn production before subsequently making their location decision based on other factors, and that their study only models this second stage location decision conditional on firms already choosing a region with a lot of corn production.

Factors that affect firm entry and location decisions considered in the previous literature include competition effects (Seim, 2006), spatial competition (Durham, Sexton and Song, 1996; Biscaia and Mota, 2013; Sesmero, Balagtas and Pratt, 2015; Wang et al., 2018), agglomeration effects (Goetz, 1997; Ellison and Glaeser, 1999; Zhu et al., 2011; Ahlfeldt et al., 2015; Gaubert, 2018; Michael Pflüger and Tabuchi, 2019; Verstraten, Verweij and Zwaneveld, forthcoming), and economies of scale (Jia, 2008).

2.2 Ethanol investment and government policy

A second strand of literature upon which we build is that on ethanol investment and the effects of government policy. The previous literature on ethanol investment includes studies that estimate the viability of ethanol plants. Many of these studies have focused largely on break-even or net present value analysis, return on investment, or similar assessments in a deterministic framework, with sensitivity analyses conducted on important costs, technologies, or prices (Whims, 2002; Gallagher et al., 2006; Eidman, 2007; Ellinger, 2007; Dal-Mas et al., 2011). To evaluate the viability of ethanol plants under stochastic conditions, price risk and cost risk have been incorporated by some studies to evaluate the profitability of a representative ethanol plant (Richardson et al., 2007; Richardson, Lemmer and Outlaw, 2007; Gallagher, Shapouri and Brubaker, 2007; Dal-Mas et al., 2011); in addition, Jouvet, Le Cadre and Orset (2012) also incorporate uncertainty in demand and competition. Markel, Sims and English (2018) use a real options framework to isolate the effect of fuel market uncertainty and policy uncertainty on the decision to enter and exit the biofuel market.

Other studies of ethanol investment have estimated the most profitable plant size under different market conditions (Gallagher, Brubaker and Shapouri, 2005; Gallagher, Shapouri and Brubaker, 2007; Khoshnoud, 2012). Several recent studies analyze ethanol plant investment option values (Schmit, Luo and Tauer, 2009; Gonzalez, Karali and Wetzstein, 2012) based on engineering cost information and various simulations.

The previous literature also studies of how government policies impact investment in ethanol plants. Schmit, Luo and Tauer (2009) and Schmit, Luo and Conrad (2011) use dynamic programming methods to show that without government policies, the recent expansionary periods would have not existed and market conditions in the late 1990s would have led to some plant closure. Babcock (2013) similarly finds that government support is important for the development of ethanol industry. On the other hand, Babcock (2011) argues that the recent high gasoline prices and phase-out of MTBE increased ethanol prices far above levels needed to justify investment in a corn ethanol plant, which means that government support might not be necessary. Cotti and Skidmore (2010) find that state-level producer tax credits can have a significant effect on a state's ethanol production capacity. Other studies have examined the effect of government policies on investment in ethanol plants econometrically (Herath Mudiyanselage, Lin and Yi, 2013; Thome and Lin Lawell, 2018; Yi and Lin Lawell, 2018a; Yi and Lin Lawell, 2018b; Yi, Lin Lawell and Thome, 2018). The previous literature also includes studies of the Renewable Fuel Standard and the effects of renewable fuel mandates on markets and/or welfare (de Gorter and Just, 2009; Lapan and Moschini, 2012; Holland et al., 2014; Chen et al., 2014; Skolrud et al., 2016; Lemoine, 2016; Moschini, Lapan and Kim, 2017; Just, 2017; Skolrud and Galinato, 2017; Korting and Just, 2017; Thome and Lin Lawell, 2018; Lade, Lin Lawell and Smith, 2018a; Lade, Lin Lawell and Smith, 2018b; Irwin, McCormack and Stock, 2018; Korting, de Gorter and Just, forthcoming).

3 Empirical Model

To analyze how strategic interactions, government policy, and economic factors affect ethanol plant entry decisions over the period 1996-2008, we estimate a discrete response panel model.

Because we are modeling the decision to invest in building a new ethanol plant, we use the terms 'investor' and 'entrant' interchangeably.

We regress the probability of ethanol plant entry on the covariates using the following logit fixed effects model:

$$Pr(I_{ikt} = 1) = 1 - F(-(N'_{kt}\delta_N + G'_{kt}\delta_G + X'_{kt}\delta_X + Year_t{}'\gamma + \nu_k)), \tag{1}$$

where I_{ikt} is an indicator of whether potential entrant i enters by building a new ethanol plant in county k in year t, and $F(\cdot)$ is the logistic cumulative distribution function.

We also estimate the following linear probability fixed effects model:

$$Pr(I_{ikt} = 1) = N'_{kt}\delta_N + G'_{kt}\delta_G + X'_{kt}\delta_X + Year_t{}'\gamma + \nu_k. \tag{2}$$

There are several reasons to estimate both the fixed effects logit and fixed effects linear probability models. The linear probabiliy model is easier to implement and the estimates are consistent if we control for the heteroskedastic errors. The logit model in (1) is preferred, however. In this particular dataset, there are relatively few instances of ethanol plant entry. Because the probability of entry is relatively low, we are on the left side of the distribution, and it is therefore advantageous to use the logit model.

In both discrete response models, $Year_t$ is either a year effect or a time trend depending on the specification. ν_k is the county fixed effect that controls for unobservable county traits, such as size or promotion of business development, which remain fixed over time. The errors for the fixed effect logit are calculated with the observed information matrix. The errors for the linear probability model are clustered at the county level.

The coefficient on the number of existing plants N_{kt} measures the net effects of the competition and agglomeration effects. Since N_{kt} is the number of existing ethanol plants open at the start of the period in which the entry decision is made, and is therefore predetermined before the entry decision is made, it is not endogenous. In the data, the maximum number of plants in existence in any county is three. Because of the time necessary to construct a plant, the potential entrant necessarily observes previously existing plants before deciding whether to enter. We include two measures of the number of existing plants N_{kt}, the number of existing plants in the county (*existing plants*), and the number of existing plants in the counties bordering the given county (*spatial lag of existing plants*). The coefficients on *existing plants* and *spatial lag of existing plants* tell us the impact of existing competitor plants on the probability of entry.

The covariates in G_{kt} describe the policy environment faced by the corn-ethanol industry. State and federal policies can affect the expected payoff from entering through the cost of inputs, expected revenues, and building costs. At the federal level we include indicators for the two versions of the Renewable Fuel Standard (RFS1 and RFS2), which are implemented as blending mandates. At the state-level, we include the year the MTBE ban was implemented; MTBE was a popular oxygenate used to meet environmental regulations and also to boost octane level, and ethanol is a substitute for MTBE. We also include state-level production tax credits.

In particular, the vector G_{kt} contains indicators of the different policies; the state policies *MTBE Ban* and *Tax Credit* are used in all specifications, and *RFS1*, *RFS2* are included in specifications with continuous $Year_t$. X_{kt} contains the following exogenous covariates: *corn price*, *soy price*, *corn intensity* and its spatial lag, *soy intensity*, *cow density*, *electric price*, *natural gas price*, *gasoline price*, *ethanol price*, and the indicator *biodiesel plant*. Like the RFS variables, we can only identify *ethanol price* in the specifications without individual year effects because there is no spatial variation in the ethanol price in our data.

The covariates in X_{kt} include economic factors that affect the payoffs from entering. On the revenue side, we include ethanol price; gasoline price; and proximity to cattle, which is a proxy for sales price of distillers' grains (DDGS, or distillers' dried grains with solubles, is a co-product of corn-ethanol production which is used for animal feed).[2] Gasoline price

[2]The co-product market is becoming more significant due to lower prices for ethanol (Dhuyvetter, Kastens and Boland, 2005). There is significant variability in participation in co-product markets (Perrin, Fretes and

could have a positive or negative impact on entry depending on whether ethanol is viewed as as an energy substitute for gasoline or an oxygenate (additive), respectively.

The vector X_{kt} also includes covariates describing the cost of ethanol production. One important factor is availability and cost of corn, the primary feedstock in the region of focus; local availability is important because transportation is costly (USDA, 2007). Corn is the largest variable cost in ethanol production (Kwiatkowski et al., 2006; Perrin, Fretes and Sesmero, 2009). In addition to corn availability and price, X_{kt} also includes soy availability and price; these variables help describe the intensity of a county's corn production. We include the natural gas price because it is a major energy source for milling corn. We also include electricity price; electricity is an important energy source in some plants.

We also control for the existence of a biodiesel plant because biodiesel and ethanol plants may compete indirectly in the feedstock market: while biodiesel plants use soy as a feedstock, much of the Midwest can be planted to soy or corn. Also, an ethanol plant may be built to satisfy a community need for crop value-added, and a biodiesel plant may compete for support.

We do not explicitly model transportation costs because data on transportation costs infrastructure is generally time-invariant, which means the impact cannot be identified as these variables are absorbed by the county fixed effects.[3] However, we do include a metro area indicator, which could capture proximity to market, as well as the potential costs of regulations.

Sesmero, 2009). Participation is driven by mill type and plant age; wet mills (corn syrup) and dry mills (DDGS) produce different co-products (DOE, 2008).

[3]In cases where the transportation infrastructure is not time-invariant, then it is likely to be endogenous at the county level. The modeling of transportation infrastructure investment decisions, which has been studied elsewhere (Fatal et al., 2012), is beyond the scope of this paper.

4 Data

4.1 Time Frame and Focus Region

We focus on corn-ethanol plants in the Midwestern United States over the period 1996 to 2008. While ethanol is produced throughout the United States using various feedstocks, 95% of the ethanol produced in this time frame is produced from corn. Focusing on corn-ethanol plants eliminates the need to consider feedstock choice in the model.[4] The majority of corn (and ethanol from corn) is produced in the Midwestern United States, so we focus on ethanol plant entry in this region, specifically in the following ten states: Iowa, Illinois, Indiana, Kansas, Minnesota, Missouri, Nebraska, Ohio, South Dakota, and Wisconsin.

We focus on the time period 1996 to 2008, which corresponds to the latest ethanol boom in the US. This time period is narrow enough to allow us to use one set of policy variables, as well as ensure similarity in plant technology. Starting the analysis earlier would also be difficult because plant startup and closure information is not readily available before this date.[5] Figure 1 shows the number of ethanol plants at the beginning and end of our study period.

Though the start-up month for new plants is available, we use annual observations for three reasons. First, the feedstock of focus, corn, has one growing season in the US. Second, construction of an ethanol plant takes significantly longer than a month, but usually less than a year, from the start of physical construction activities.[6] Finally, much of the data on

[4]For structural econometric models of feedstock choice, see Yi and Lin Lawell (2018b), who model ethanol investment and feedstock choice in Europe; and Yi and Lin Lawell (2018a), who model ethanol investment and feedstock choice in Canada.

[5]Including the entrants during 2009 and 2010 would require accounting for plant closure due to the market crash and implosion of Verasun, a large producer. Many plants stopped production in late 2008 or early 2009 following Verasun's bankruptcy declaration on October 31, 2008. Operations were normal the rest of the year, and many of the shuttered plants have since restarted under new ownership. Prior to 2008, there was only one permanent closure (exit) in the sample; others closures were the result of accidents or buyouts, and the plants returned to normal operations. The exit phenomenon in a subject of ongoing work and is outside the scope of this model.

[6]There was a production bottleneck in 2007, when plants took 18-24 months to build (Koplow, 2007). We do not consider announcements of new plants, as other studies did, because many announced projects were never completed as investors fell through before construction began.

other variables are publicly available at an annual level.

We eliminate completely non-agricultural counties within the ten states (e.g. northern Minnesota), as well as those with missing data on agricultural production, resulting in a sample with 855 unique counties. This results in potentially 11,115 county-year observations over the thirteen-year time period. We add another dimension to account for the number of potential entrants in each county-year.

4.2 Plant Variables

Our ethanol plant data set includes information about start-up date of new entrants, and nameplate capacity and ownership type for new and existing plants. 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.

The sample begins with 22 operational plants at the start of 1996, and ends with 149 operational plants with a total capacity of almost 10 billion gallons per year in 2008. Figure 1 maps the number of operational ethanol plants by county in the first and last years of our data set, respectively.

The entry variable I_{ikt} is an indicator of whether potential entrant i enters by building a new ethanol plant in county k in year t. As the maximum number of ethanol plants in any county in our data set during the time period of our data set is three, we allow for up to 3 potential ethanol plant entrants per county-year. The entry variable I_{ikt} is equal to 1 if the plant enters in a given calendar year.[7] Once a potential entrant i enters, it is no longer a potential entrant and therefore exits the sample.

The number of existing plants N_{kt} in the county measures the number of operational

[7]Entry is the date of the first grind of corn, which is the first step in corn-ethanol production.

plants in that county on January 1 of year t, and is therefore observable to any potential entrant making a decision in year t. In an alternate specification, we define N_{kt} as a continuous variable of capacity of existing plants.[8] We also define a spatial lag of the existing plant variable as the number of plants in the counties bordering a given county.

The dataset on biodiesel plants was constructed in the same manner as the ethanol plant variables. The original biodiesel plant lists were from the National Biodiesel Board and Biodiesel Magazine. Analogous to the number of existing plants, we construct an indicator variable that signals the existence of a biodiesel plant in county k at the start of the calendar year t.

4.3 Policy Variables

We include state-level policy variables. The first state-level policy variable we use is an indicator of whether the state banned MTBE at any point in a given year. All the Midwestern states in our sample implemented MTBE bans by 2005, before the nationwide ban took effect in 2006.

The second state-level policy variable represents the state producer tax credits.[9] Defining this variable is complicated because each state places different contingencies on receiving these funds. For example, some states support only large-capacity plants, others only small or community-owned plants. Thus, even in states with tax credits, not all entering or incumbent plants qualify. In addition, some of the credits are available for a specified number or years, while others expire on a date unrelated to time of plant entry. Because of these differences, we represent these policies with a binary variable indicating if producer tax credit benefits were offered to plants that entered in that year, and test the robustness to that specification.

For federal-level policy variables, we specify two variables to capture the effects of

[8] Capacity is a good proxy for production because plants operate continuously at or near nameplate, except during regular maintenance (Kwiatkowski et al., 2006).

[9] The American Coalition for Ethanol (2007) provides detailed description and review of the policies. Cotti and Skidmore (2010) study state-level impacts of these policies.

the Renewable Fuel Standards (RFS).[10] The RFS was created under the Energy Policy Act of 2005 with the goal of accelerating the use of fuels derived from renewable sources (EPA, 2013). This initial RFS (RFS1) mandated that a minimum of 4 billion gallons of ethanol be blended into gasoline in 2006, rising to 7.5 billion gallons by 2012. Two years later, the Energy Independence and Security Act of 2007 greatly expanded the biofuel mandate volumes, creating the RFS2. The RFS2 requires steadily increasing volumes of biofuel to be blended into the nation's fuel supply, reaching 37 billion gallons a year by 2022. We model RFS1 with an indicator for the years 2005 and 2006 and RFS2 as an indicator for the years 2007 and 2008.

4.4 Other Data

Corn and soy prices are available annually from the National Agricultural Statistics Service of the USDA (NASS) at the state level. Corn and soy production and acreage are available annually by county from NASS. Because counties are different areas, we construct a county corn intensity variable, defined as the corn acreage divided by the total area of the county, to capture area-independent acreage using county acreage from the US Census.[11] We also construct a spatial lag of the corn intensity variable as well as a county-level soy intensity variable. Because corn price data are not publicly available at a county level, the local competition in the corn feedstock market is captured both by the county-level corn intensity variable and by the covariate N_{kt} measuring the number of existing plants in the county.

To represent the potential market for distillers' grains (DDGS), a co-product of corn-ethanol production that is used for animal feed, we construct a county-level cow density variable using the number of cows per district-acre, where the number of cows is the count of 'all cattle', available from NASS, and districts are defined by the USDA.[12] The potential DDGS market also includes hogs, but data is not available at the district level for all states.

[10] We do not include other federal-level policy variables such as tax credit or the small producer subsidy in the analysis because they do not vary enough in the time period to identify the effects.

[11] As a robustness test, we also run specifications defining corn intensity as production over area.

[12] A district is made of up to 120 counties and there are usually 6-8 districts per state.

However, because cattle use DDGS more efficiently than hogs, they represent the larger market for co-products (NASS, 2007).

The ethanol price is the free on board price in Omaha, and is published by the Nebraska Energy Office. We use state-level total gasoline rack prices from the Energy Information Administration. We do not include an E85 price in this analysis because the price series began much more recently than our time frame, and it lacks spatial variation. Natural gas (city gate) price and electricity price to industry are available annually from the EIA, also at state level.[13] We use the average urban CPI to deflate all the prices. The final variable, an indicator for metropolitan areas, is the US Census definition of counties in metropolitan statistical areas.

Because we do not have local variation in ethanol, gasoline, natural gas, or electricity prices, local competition in the ethanol and gasoline output markets and in the gasoline, natural gas and electricity input markets are captured by the covariate N_{kt} measuring the number of existing plants in the county.

The summary statistics for the explanatory variables used in our empirical analysis are presented in Table 1.

5 Results

The results from the estimation of the fixed effects logit model in equation (1) and the linear probability fixed effects model in equation (2) are presented in Table 2.

The results of the logit estimation in (1) are in Specifications A and B of Table 2. The coefficient on *existing plants* is large, negative, and significant; indicating that the number of existing plants has a negative effect on the probability of entry in a given county. The coefficient on the *spatial lag of existing plants*, the number of plants in the counties bordering a given county, is positive, though insignificant, indicating a potential positive agglomeration

[13]We use city gate natural gas price instead of price to industry because the complete series is available; these two price series trend together within a given state.

effect regionally. The significant negative sign on *existing plants* that is only present within a county confirms the existence of localized competition also seen in Sarmiento, Wilson and Dahl (2012) and Lambert et al. (2008).

One reason we see so few significant variables in this regression is because the fixed effects logit relies upon within-county variation for identification, meaning we estimate (1) for only the counties that had an entrant in the time period. Variables such as *corn intensity* vary more spatially than they do across time, hence this regression does not detect an impact on probability of entry. Goetz (1997) suggests a two-stage location-selection process where firms chose a region based on some factors, and then enter in a specific location based on others as in Haddad, Taylor and Owusu's (2010) regional model of ethanol plant location. Corn-ethanol plants are location in regions with high corn availability, but variation within the region (and in our case, over time) is not large. Instead, local and market factors drive location and entry decisions.

The magnitude, and sometimes sign, of some of the other coefficients depends on the specification of time $Year_t$ in the regression model. A continuous $Year_t$ controls for changes in technologies and preferences over time, while the individual $Year_t$ effects also capture events, policies, market conditions at the national level. The coefficients on *natural gas price*, *corn price*, and *soy price* change sign and magnitude across the regression, though none are significantly different from zero. These variables are all correlated and trend upwards over time, which may make their effects difficult to distinguish from the time trend.

The coefficient on *Tax Credit* is positive but insignificant in the regression with year effects (Specification A of Table 2), and is larger and becomes significant in the regression with continuous time (Specification B). We find no significant impact of *RFS1*, *RFS2*, or the *MTBE Ban*. Cotti and Skidmore (2010) found positive impacts of state ethanol tax credits on state ethanol capacity, suggesting that these policies drive regional location choices, but that perhaps national-level policies drive overall growth in the industry.

The only other significant coefficient is that on *gasoline price*, which also has a large, positive effect on the probability of entry, indicating that potential entrants may view ethanol as a gasoline substitute. We explore this result further below.

The linear probability fixed effects model estimation of equation (2) is interesting for two reasons. First, the results in Specifications C and D of Table 2 serve as a comparison to the logit fixed effects model in (1). The signs and significance levels of the linear probability model estimates in Specifications C and D are qualitatively similar to the logit estimates. Like the logit model, the linear probability models in Specifications C and D are estimated only for the counties k that have an entrant at some point in the period.

A second reason the linear probability model is informative is that we can include the full data sample and account for within and cross-sectional variation. As seen in Specifications E and F of Table 2, we find more significant variables for the linear probability model when we use the full dataset and not just the counties who have entrants. While *existing plants* still have a negative and significant effect on entry, we see that the effect of *spatial lag of existing plants* is positive and significant. These results suggest that the net negative competitive effect among plants not only dissipates with distance, but also becomes net positive, indicating possible agglomeration benefits in the ethanol industry.

We run several specification tests of the discrete response models in equations (1) and (2). Their results are presented in Table 3. First, we use a Hausman test to choose between random effects and fixed effects. The Hausman χ^2 statistics from the test on the restricted and full random effects models are all very large, with corresponding p-values of 0.000, indicating county unobservables are likely to be correlated with the regressors, and therefore that fixed effects is the appropriate specification.[14]

We test for potential endogeneity of *corn intensity* using a Durbin-Wu-Hausman test. In the first-stage regression, the instruments for *corn intensity* are the time lags of *corn intensity* and *corn price*. The estimated coefficient on the first-stage residuals in the

[14] The restricted random effects model includes the same regressors as the fixed effects model, while the full model includes the time-invariant regressors, allowing accounting for potential efficiency gain from their inclusion (Wooldridge, 2010).

second stage regression is insignificant, indicating that we cannot reject the exogeneity of corn intensity in any specification.[15]

We do not anticipate endogeneity problems with the other variables such as *corn price* because 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. For example, McNew and Griffith (2005) find that while ethanol plants increase the basis for corn price, this effect is limited to around 50 miles from the plant, while the price variables in this analysis are measured at the state level. An additional argument for using contemporaneous prices rather than futures prices in our model is that while futures prices exist, they are at a national level, and therefore will be absorbed by the year effects.

In Table A.1 in Appendix A, we estimate the logit model (1) with alternate specifications of the corn and soy variables (*corn price*, *soy price*, *corn intensity*, *soy intensity*). We construct ratios of *corn to soy price* and *corn to soy intensity* and include them in the regressions in place of, and as well as, the previously specified variables. The hypothesis is that perhaps the relative prices and production intensities may capture more variation in entry probability than the levels. However, the results are not qualitatively different from the results in Table 2.

In Table A.2 in Appendix A, we explore the large positive effect of *gasoline price* further by estimating the logit model (1) with alternate specifications of the *ethanol price* and *gasoline price* variables. We construct a ratio of the *ethanol to gasoline price* and include in the regression in place of, and as well as, the individual price variables. One advantage of this alternate specification is that we can control for the ethanol price regardless of the specification of $Year_t$. While *ethanol price* is measured at the national level, the ethanol-gasoline price ratio is at the state level.

The alternate specification of *gasoline price* and *ethanol price* does not have any

[15] As a robustness check, we estimate the models with a time lagged corn intensity variable instead of contemporaneous corn intensity. There is no significant difference in the other estimates (results not reported).

qualitative effects on other coefficient estimates, except for the coefficients on *RFS1* and *RFS2*. The estimates of the RFS impacts are larger, and significant, when *ethanol to gasoline price ratio* is included in the regression. Additionally, we detect a positive impact of *ethanol price* in the specifications with continuous $Year_t$. In these specifications, the coefficient on *ethanol to gasoline price* is small and insignificant. In the specifications with year effects instead of a continuous $Year_t$, the effect of *ethanol to gasoline price* on the probability of entry is large, negative, and significant, which supports the view of ethanol and gasoline as substitutes. Babcock (2012) discusses the relative cost of gasoline and ethanol in a policy context, and finds market scenarios in which ethanol can be viewed as an energy substitute for gasoline, and others in which ethanol is viewed as an additive.

In Specifications T and U of Table A.3 in Appendix A, we estimate the logit model (1) with an alternate specification for *Tax Credit*: we model the effect of the expected lifetime value of the tax credit instead of an indicator for the existence of the policy. In Specifications V and W of Table A.3, we estimate the logit model (1) with an alternate specification for *existing plants*: we model capacity instead of count of other plants. In all cases, there are no qualitative differences in the results.

To allow for the potential for differing impacts of competitors based on the plant ownership type of the existing plants, we group the existing plants by ownership type: singlets, ethanol-focused firms, and conglomerates. The number of existing plants in each county and region by ownership type are presented in Table 4. Different types of operators may produce different externalities (either positive and negative) towards potential entrants due to different linkages to related markets and/or the community.

Table 5 presents the results for 3 groupings of existing plants by ownership, as well as the results for Specification A from Table 2 for comparison. Specification X disaggregates the number of existing plants by singlets versus non-singlets. Specification Y disaggregates the number of existing plants by conglomerates versus non-conglomerates. Specification Z disaggregates the number of existing plants by singlets, ethanol-focused firms and conglomerates. All specifications in Table 5 include the same policy variables and economic variables as Specification A.

According to the results in Table 5, existing plants in the county have the same negative and significant impact on entry as in Specification A, regardless of ownership. In contrast, the *spatial lag of existing plants*, the number of plants in the counties bordering a given county, has either an insignificant or positive effect on entry. In particular, nearby conglomerates and large ethanol producing firms have a negative effect on entry while those a bit farther away in neighboring counties have a positive effect. For the singlet plants, the effect of existing singlet plants is significant and negative at the county level and insignificant at the regional level. The negative county affect from all types of plants indicates they may be competing for corn as an input; we would expect this effect to decline or even disappear at the region level following the results of McNew and Griffith (2005) that most corn in sourced within 50 miles of the plant, well within the average size of a county.

The fact that only some types of plants in neighboring counties have a positive effect indicates the potential for agglomeration externalities. Larger or conglomerate plants may have linkages to other markets and well-developed infrastructure for shipping ethanol or a well-trained workforce may be useful to plants entering in a neighboring county. Singlet plants may not provide this sort of benefit because they are generally smaller (have fewer employees, etc.), and also, many (though not all) were developed as a local value-added source, so they may not have developed other linkages.

When controlling for the ownership type of the existing ethanol plants, the estimates of the coefficients on the policy variables G_{kt} and economic variables X_{kt} are very close to the estimates from the Table 2 (results not reported).

6 Conclusions

In this paper we examine how strategic interactions, government policy, and economic factors affect ethanol plant entry decisions at the county level using discrete response models. We focus on corn-ethanol plants in the Midwestern United States, where the majority of corn in the US is grown, over the period 1996-2008.

Results show that that existing plants have a negative effect on the probability of entry in a given county. We also find evidence that existing plants in the counties bordering a given county may have a positive effect on entry in that county. These results suggest that the net negative competitive effect among plants not only dissipates with distance, but also becomes net positive, indicating possible agglomeration benefits in the ethanol industry.

We find in particular that existing conglomerates and large ethanol producing firms in neighboring counties have a positive effect on ethanol plant entry, while existing singlet plants in neighboring counties do not. Larger or conglomerate plants may have linkages to other markets and well-developed infrastructure for shipping ethanol or a well-trained workforce may be useful to plants entering in a neighboring county. Singlet plants may not provide this sort of benefit because they are generally smaller (have fewer employees, etc.), and also, many (though not all) were developed as a local value-added source, so they may not have developed other linkages.

Our results indicate that there is an important strategic component to ethanol plant entry decisions. This net negative effect of existing plants in a county may be due to localized competition. Results also show that the ownership type of the existing plant matters: nearby conglomerates and large ethanol producing firms have a negative effect on entry while those a bit farther away (in the same region, but not the same county) have a positive effect. This result is consistent with previous estimates that the competition for feedstock is local (McNew and Griffith, 2005).

In the previous literature, Sarmiento, Wilson and Dahl (2012) find that the negative competition effect decays with distance. Their empirical specification does not allow for the possibility of a positive agglomeration effect.

Our result is a refinement of the previous literature because not only do we find a net-negative local externality from competition, but we also detect a positive externality when the existing neighboring plant is outside the source area for feedstock.

Our results also indicate that the intensity of corn production is an important determinant of ethanol plant entry. Approximately 70% of the cost of producing ethanol is the feedstock cost, of which the transportation costs for bulky grains constitute a significant share (Whittington, 2006). As a consequence, the distance from a plant to the feedstock production area is extremely important, and local production intensity of feedstock matters.

References

Ahlfeldt, G., S. Redding, D. Sturm, and N. Wolf. (2015). The economics of density: Evidence from the Berlin Wall. *Econometrica*, 83 (6), 21272189.

American Coalition for Ethanol. (2007). *STATUS: A state by state handbook*. Sioux Falls, SD.

Anderson, S., and A. Elzinga. (2014). A ban on one is a boon for the other: Strict gasoline content rules and implicit ethanol blending mandates. *Journal of Environmental Economics and Management*, 67 (3), 258-273.

Babcock, B.A. (2011). The impact of U.S. biofuel policies on agricultural price levels and volatility. International Centre for Trade and Sustainable Development Issue Paper 35.

Babcock, B.A. (2013). Ethanol without subsidies: An oxymoron or the new reality? *American Journal of Agricultural Economics*, 95 (5), 1317-1324.

Bartik, T.J. (1985). Business location decisions in the United States: Estimates of the effects of unionization, taxes, and other characteristics of states. *Journal of Business and Economic Statistics*, 3 (1), 14-22.

Behrens, K., W.M. Brown, and T. Bougna. (2018). The world is not yet flat: Transport costs matter! *Review of Economics and Statistics*, 100 (4), 712-724.

Biodiesel Magazine. (2008). Biodiesel Plant Lists. Last accessed online July 2008. URL: `biodieselmagazine.com`.

Biscaia, R., and I. Mota. (2013). Models of spatial competition: A critical review. Papers in Regional Science, 92 (4), 851-871.

Bushnell, J.B., J.E. Hughes, and A. Smith. (2017). Food vs. fuel?: Impacts of petroleum shipments on agricultural prices. NBER Working Paper No. 23924.

Carter, C.A., G.C. Rausser, and A. Smith. (2011). Commodity booms and busts. *Annual Review of Resource Economics*, 3, 87-118.

Chen, X., H. Huang, M. Khanna, and H. Onal. (2014). Alternative transportation fuel standards: Welfare effects and climate benefits. *Journal of Environmental Economics* *and Management*, 67 (3), 241-257.

Coase, R. (1960). The problem of social cost. *Journal of Law and Economics*, 3, 1-44.

Cotti, C., and M. Skidmore. (2010). The impacts of state government subsidies and tax credits in an emerging industry: Ethanol production 1980-2007. *Southern Economic Journal*, 76 (4), 1076-1093.

Dal-Mas, M., S. Giarola, A. Zamboni, and F. Bezzo. (2011). Strategic design and investment capacity planning of the ethanol supply chain under price uncertainty. *Biomass and Bioenergy*, 35 (5), 2059-2071.

de Gorter, H., D. Drabik, and D.R. Just. (2013). How biofuels policies affect the level of grains and oilseed prices: Theory, models and evidence. *Global Food Security*, 2, 82-88.

de Gorter, H., D. Drabik, D.R. Just, and E.M. Kliauga. (2013). The impact of OECD biofuels policies on developing countries. *Agricultural Economics*, 44, 477-486.

de Gorter, H., and D.R. Just. (2009). The economics of a blend mandate for biofuels. *American Journal of Agricultural Economics*, 91 (3), 738-750.

de Gorter, H., and D.R. Just. (2010). The social costs and benefits of biofuels: The intersection of environmental, energy and agricultural policy. *Applied Economics Perspectives and Policy*, 32 (1), 4-32.

Dhuyvetter, K.C., T.L. Kastens, and 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.

DOE [Department of Energy]. (2008). Energy Time Lines: Ethanol. Revised June 2008. Washington, DC: DOE.

Durham, C.A., R.J. Sexton, and J.H. Song. (1996). Spatial competition, uniform pricing, and transportation efficiency in the California processing tomato industry. American Journal of Agricultural Economics, 78 (1), 115-125.

EIA [Energy Information Administration of the DOE]. (2009). Total Gasoline Rack Price, Total Industry Industrial Electrical Price, Natural Gas City Gate Price. Last accessed online July 2009. URL http://tonto.eia.doe.gov/dnav.

EIA [Energy Information Administration of the DOE]. (2015). New EIA Monthly Data Track Crude Oil Movements by Rail. Last accessed online April 2015. URL: http://www.eia.gov/todayinenergy/detail.cfm?id=20592#tabs_Slider-5.

Eidman, V.R. (2007). Ethanol economics of dry mill plants. *Corn-Based Ethanol in Illinois and the US: A Report from the Department of Agricultural and Consumer Economics* (pp. 22-36). University of Illinois.

Ellinger, P.N. (2007). Assessing the financial performance and returns of ethanol production: a case study analysis. *Corn-Based Ethanol in Illinois and the US: A Report from the Department of Agricultural and Consumer Economics* pp. 37-62. University of Illinois.

Ellison, G., and E.L. Glaeser (1999). The geographic concentration of industry: Does natural advantage explain agglomeration? *American Economic Review*, 89 (2), 311-316.

Ethanol Producer Magazine. (2010). Plant lists. Accessed online December 2010. URL: www.ethanolproducer.com.

Fatal, Y.S., S. Kotisir, H.A. Tejeda, and C. Zhan. (2012). Reducing GHG emissions and energy input in the U.S. supply chain of ethanol and gasoline. Working paper, North Carolina State University.

Gallagher, P. (2009). Roles for evolving markets, policies, and technology improvements in U.S. Corn Ethanol Industry Development. *Regional Economic Development*, 5 (1), 12-33.

Gallagher, P.W., H. Brubaker, and H. Shapouri. (2005). Plant size: Capital cost relationships in the dry mill ethanol industry. *Biomass and Bioenergy*, 28 (6), 565-571.

Gallagher, P., G. Schamel, H. Shapouri, and H. Brubaker. (2006). The international competitiveness of the US corn-ethanol industry: A comparison with sugar-ethanol processing in Brazil. *Agribusiness*, 22 (1), 109-134.

Gallagher, P., H. Shapouri, and H. Brubaker. (2007). Scale, organization, and profitability of ethanol processing. *Canadian Journal of Agricultural Economics/Revue canadienne* *d'agroeconomie*, 55 (1), 63-81.

Gaubert, C. (2018). Firm sorting and agglomeration. *American Economic Review*, 108 (11), 3117-3153.

Goetz, S. (1997). State- and county-level determinants of food manufacturing establishment growth: 1987-93. *American Journal of Agricultural Economics*, 79, 838-850.

Gonzalez, A.O., B. Karali, and M.E. Wetzstein. (2012). A public policy aid for bioenergy investment: Case study of failed plants. *Energy Policy*, 51, 465-473.

Haddad, M. A., G. Taylor, and F. Owusu. (2010). Locational choices of the ethanol industry in the Midwest corn belt. *Economic Development Quarterly*, 24 (1), 74-86. Helfand, G. (1992). Standards versus standards: The effects of different pollution restrictions. *American Economic Review*, 81, 622-634.

Herath Mudiyanselage, N., C.-Y.C. Lin, and F. Yi. (2013). An analysis of ethanol investment decisions in Thailand. *Theoretical Economics Letters*, 3 (5A1), 14-20.

Hochman, G., S.E. Sexton, and D.D. Zilberman. (2008). The economics of biofuel policy and biotechnology. *Journal of Agricultural and Food Industrial Organization*, 6 (2), Article 8.

Holland, S., C. Knittel, and J. Hughes. (2009). Greenhouse gas reductions under low carbon fuel standards? *American Economic Journal: Economic Policy*, 1 (1), 106-146.

Holland, S., C. Knittel, J. Hughes, and N. Parker (2014). Some inconvenient truths about climate change policies: The distributional impacts of transportation policies. *Review of Economics and Statistics*, 97 (5), 1052-1069.

Irwin, S.H., K. McCormack, and J.H. Stock. (2018). The price of biodiesel RINs and economic fundamentals. NBER Working Paper No. 25341.

Jia, P. (2008). What happens when Wal-Mart comes to town: An empirical analysis of the discount retail industry. Econometrica, 76 (6), 1263-1316.

Jouvet, P.-A., E. Le Cadre, and C. Orset. (2012). Irreversible investment, uncertainty, and ambiguity: The case of bioenergy sector. *Energy Economics*, 34 (1), 45-53.

Just, D. R. (2017). Comment on 'The Renewable Fuel Standard in competitive equilibrium: Market and welfare effects'. *American Journal of Agricultural Economics*, 99 (5), 1143-1145.

Khoshnoud, M. (2012). Quantity and Capacity Expansion Decisions for Ethanol in Nebraska and a Medium Sized Plant. Master thesis, University of Nebraska-Lincoln.

Koplow, D. (2007). Biofuels- At what cost? Government support for ethanol and biodiesel in the United States: 2007 Update. Prepared for Global Subsidies Initiative (GSI) of the International Institute for Sustainable Development (IISD), Geneva.

Korting, C., H. de Gorter, and D.R. Just. (forthcoming). Who will pay for increasing biofuel mandates?: Incidence of the Renewable Fuel Standard given a binding blend wall. *American Journal of Agricultural Economics*.

Korting, C., and D.R. Just, D.R. (2017). Demystifying RINs: A partial equilibrium model of U.S. biofuel markets. *Energy Economics*, 64, 353-362.

Kwiatkowski, J.R., A.J. McAloon, F. Taylor, and D.B. Johnston. (2006). Modeling the process and costs of fuel ethanol production by the corn dry-grind process. *Industrial Crops and Products*, 23, 288-296.

Lade, G.E., and C.-Y.C. Lin Lawell. (2015). The design and economics of low carbon fuel standards. *Research in Transportation Economics*, 52, 91-99.

Lade, G.E., and C.-Y.C. Lin Lawell. (2018). The design of renewable fuel policies and cost containment mechanisms. Working paper, Cornell University.

Lade, G.E., C.-Y.C. Lin Lawell, and A. Smith. (2018a). Designing climate policy: Lessons from the Renewable Fuel Standard and the blend wall. *American Journal of Agricultural Economics*, 100 (2), 585-599.

Lade, G.E., C.-Y.C. Lin Lawell, and A. Smith. (2018b). Policy shocks and market-based regulations: Evidence from the Renewable Fuel Standard. *American Journal of Agricultural Economics*, 100 (3), 707-731.

Lambert, D.M., M. Wilcox, A. English, and L. Stewart. (2008). Ethanol plant location determinants and county comparative advantage. *Journal of Agricultural and Applied Economics*, 40, 117-135.

Lapan, H., and G. Moschini. (2012). Second-best biofuels policies and the welfare effects of quantity mandates and subsidies. *Journal of Environmental Economics and Management*, 63, 224-241.

Lemoine, D. (2016). Escape from third-best: Rating emissions for intensity standards. *Environmental and Resource Economics*, 67(4), 789-821.

Lin Lawell, C.-Y.C. (2017). Dynamic structural econometric modeling of the ethanol industry. In A.A. Pinto and D. Zilberman (Eds.), *Modelling, Dynamics, Optimization and Bioeconomics II* (pp. 293-306). Springer Proceedings in Mathematics & Statistics.

Markel, E., C. Sims, and B.C. English. (2018). Policy uncertainty and the optimal investment decisions of second-generation biofuel producers. *Energy Economics*, 76, 89-100.

Maxwell, C., and M. Davison. (2013). Using real option analysis to quantify ethanol policy impact on the firm's entry into and optimal operation of corn ethanol facilities. *Energy Economics*, 42, 140-151.

McNew, K., and D. Griffith. (2005). Measuring the impact of ethanol plants on local grain prices. *Review of Agricultural Economics*, 27 (2), 164-180.

Moschini, G., H. Lapan, and H. Kim (2017). The Renewable Fuel Standard in competitive equilibrium: Market and welfare effects. *American Journal of Agricultural Economics*, 99 (5), 1117-1142.

NASS [National Agricultural Statistics Service of the USDA]. (2010). Quickstats database. Last accessed December 2010. URL: `www.nass.usda.gov/QuickStats/`.

NASS [National Agricultural Statistics Service of the USDA]. (2007). Ethanol co-products used for livestock feed. Washington, DC: NASS.

National Biodiesel Board. (2008). Plant Lists. Last accessed 2008. URL `www.biodiesel.org`.

Nebraska Energy Office. (2010). Ethanol Prices. Accessed online November 2010. URL:

`www.neo.ne.gov.statshtml/66.html`.

Perrin, R.K., N.F. Fretes, and J.P. Sesmero. (2009). Efficiency in Midwest US corn ethanol plants: A plant survey. *Energy Policy*, 37 (4), 1309-1316.

Pflüger, M., and T. Tabuchi. (2019). Comparative advantage, agglomeration economies and trade costs. of Urban Economics, 109, 1-13.

Pigou, A. (1920). Economics of Welfare. Macmillan and Co.

Preonas, L. (2018). Market power in coal shipping and implications for U.S. climate policy. Working paper, University of Chicago.

Renewable Fuels Association. (2010). Plant Lists. Last accessed December 2010. URL: `www.ethanolrfa.org`.

Richardson, J.W, B.K. Herbst, J.L. Outlaw, R.C. Gill, et al. (2007). Including risk in economic feasibility analyses: The case of ethanol production in Texas. *Journal of Agribusiness*, 25 (2), 115-132.

Richardson, J.W, J.W. Lemmer, and J.L. Outlaw. (2007). Bio-ethanol production from wheat in the winter rainfall region of South Africa: A quantitative risk analysis. *International Food and Agribusiness Management Review*, 10 (2), 181-204.

Sarmiento, C., W.W. Wilson, and B. Dahl. (2012). Spatial competition and ethanol plant location decisions. *Agribusiness*, 28 (3), 260-273.

Schmit, T.M., J. Luo, and J.M. Conrad. (2011). Estimating the influence of ethanol policy on plant investment decisions: A real options analysis with two stochastic variables. *Energy Economics*, 33, 1194-1205.

Schmit, T.M., J. Luo, and L.W. Tauer. (2009). Ethanol plant investment using net present value and real options analysis. *Biomass and Bioenergy*, 33, 1442-1451.

Seim, K. (2006). An empirical model of firm entry with endogenous product-type choices. RAND Journal of Economics, 37 (3), 619-640.

Sesmero, J.P., J.V. Balagtas, and M. Pratt. (2015). The economics of spatial competition for corn stover. Journal of Agricultural and Resource Economics, 40(3), 425441.

Skolrud, T.D., and G.I. Galinato. (2017). Welfare implications of the renewable fuel standard with an integrated tax-subsidy policy. *Energy Economics*, 62, 291-301.

Skolrud, T.D., G.I. Galinato, S.P. Galinato, C.R. Shumway, and J.K. Yoder. (2016). The Role of federal Renewable Fuel Standards and market structure on the growth of the cellulosic biofuel sector. *Energy Economics*, 58, 141-151.

Starr, A.W., and Y.C. Ho. (1969). Nonzero-sum differential games. *Journal of Optimization Theory and Applications*, 3, 184-206.

Stock, J.H. (2015). The Renewable Fuel Standard: A path forward. Report, Columbia Center on Global Energy Policy.

Stock, J.H. (2018) Reforming the Renewable Fuel Standard. Working paper, Harvard University.

Thome, K.E., and C.-Y.C. Lin Lawell. (2018). Ethanol plant investment and government policy: A dynamic structural econometric model. Working paper, Cornell University.

USDA [U.S. Department of Agriculture]. (2007). *Ethanol Transportation Backgrounder: Expansion of US corn-based ethanol from the agricultural transportation perspective.* Washington DC: USDA AMS.

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

Verstraten, P., G. Verweij, and P.J. Zwaneveld. (forthcoming). Complexities in the spatial scope of agglomeration economies. *Journal of Regional Science.*

Wang, Y., M.S. Delgado, J. Sesmero, and B.M. Gramig. (2018). Market structure and the effect of ethanol expansion on land allocation: A spatially explicit analysis. Working paper, Purdue University and University of Illinois at Urbana-Champaign.

Whims, J. (2002). Corn based ethanol costs and margins. Agricultural Marketing Resource Center Report, Dept. of Ag. Econ., Kansas State University.

Whittington, T. (2006). International Overview of Ethanol Production and Policies. Department of Agriculture and Food Western Australia Report.

Wooldridge, J.M. (2010). *Econometric Analysis of Cross Section and Panel Data*. Second Edition. Cambridge, MA: MIT Press.

Yi, F., and C.-Y.C. Lin Lawell. (2018a). Ethanol plant investment in Canada: A structural model. Working paper, Cornell University.

Yi, F., and C.-Y.C. Lin Lawell. (2018b). What factors affect the decision to invest in a fuel ethanol plant?: A structural model of the ethanol investment timing game. Working paper, Cornell University.

Yi, F., C.-Y.C. Lin Lawell, and K. Thome. (2018). A dynamic model of subsidies: Theory and application to ethanol industry. Working paper, Cornell University.

Zhu, T., V. Singh, and A. Dukes. (2011). Local competition, entry, and agglomeration. Quantitative Marketing and Economics, 9 (2), 129-154.

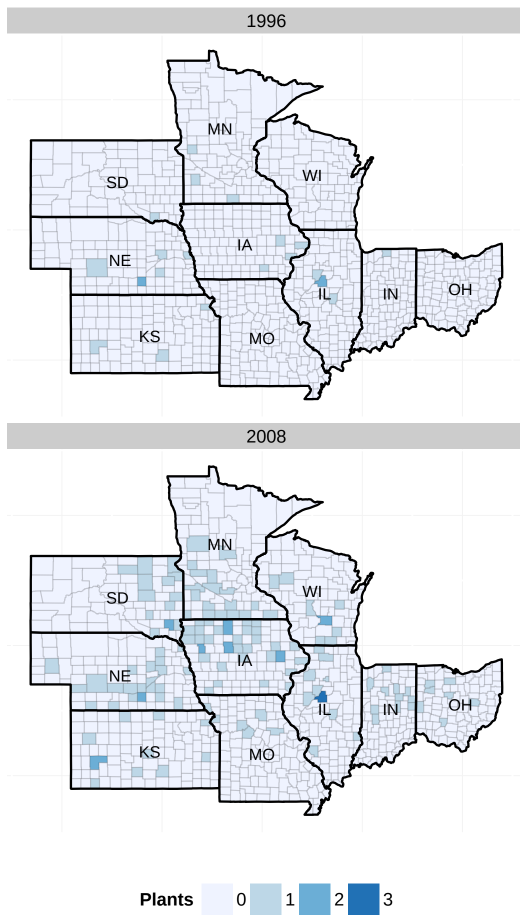

Figure 1: Number of operational ethanol plants by county in the Midwestern United States

Table 1: Summary statistics

Variable	Counties with at least one new ethanol plant (1996-2008) mean	std. dev.	Full Sample mean	std. dev.	Spatial Resolution of Data
Ethanol Plant Entry [dependent variable: indicator]	0.034	0.182	0.004	0.066	
Existing Plants [count]	0.202	0.417	0.040	0.203	county
Spatial Lag of Existing Plants [count]	0.632	1.069	0.395	0.819	contiguous bordering counties
Existing Biodiesel Plant [indicator]	0.015	0.126	0.010	0.105	county
MTBE Ban [indicator]	0.582	0.493	0.487	0.500	state
Tax Credit [indicator]	0.370	0.483	0.346	0.476	state
RFS1 [indicator]	0.167	0.373	0.163	0.369	national
RFS2 [indicator]	0.146	0.354	0.150	0.357	national
Ethanol Price [$/gallon]	1.781	0.416	1.778	0.418	national
Gasoline Price [$/gallon]	1.349	0.553	1.341	0.562	state
Natural Gas Price [$/1000 ft3]	6.527	1.852	6.516	1.917	state
Electricity Price [cents/KwH]	5.107	0.495	5.232	0.536	state
Corn Price [$/bushel]	2.785	0.675	2.828	0.673	state
Soy Price [$/bushel]	7.106	1.690	7.160	1.693	state
Corn Intensity [acres planted/total acreage]	0.299	0.130	0.200	0.144	county
Spatial Lag of Corn Intensity [acres planted/total acreage]	0.279	0.123	0.200	0.129	contiguous bordering counties
Soy Intensity [acres planted/total acreage]	0.245	0.127	0.183	0.131	county
Cow Density [head/acre]	0.103	0.055	0.085	0.051	district (USDA definition)
Number Observations	3,687		28,769		
Number of Counties	120		855		

Table 2: Results from fixed effects model of ethanol plant entry

	Dependent variable is probability of ethanol plant entry					
	Fixed Effects Logit		Linear Probability Models			
			FE Logit sample		Full sample	
	A	B	C	D	E	F
Existing Plants	-13.79***	-13.35***	-0.180***	-0.172***	-0.077***	-0.076***
	(1.722)	(1.585)	(0.011)	(0.011)	(0.004)	(0.004)
Spatial Lag of Existing Plants	0.57	0.44	0.004	0.008	0.002*	0.003**
	(0.417)	(0.378)	(0.006)	(0.006)	(0.001)	(0.001)
Existing Biodiesel Plant	-0.51	-0.10	-0.037	-0.020	-0.004	-0.003
	(1.729)	(1.450)	(0.032)	(0.032)	(0.005)	(0.005)
MTBE Ban	-0.97	-0.98	0.003	-0.006	-0.003	-0.003
	(0.927)	(0.808)	(0.013)	(0.012)	(0.002)	(0.002)
Tax Credit	0.19	1.22*	0.006	0.015	-0.001	0.000
	(0.707)	(0.585)	(0.008)	(0.008)	(0.001)	(0.001)
RFS1		0.72		-0.002		-0.002
		(1.433)		(0.020)		(0.003)
RFS2		0.50		0.058		-0.000
		(3.067)		(0.041)		(0.005)
Gasoline Price	29.42*	5.40	1.120***	0.118*	0.105***	0.013*
	(12.462)	(3.230)	(0.246)	(0.048)	(0.031)	(0.006)
Ethanol Price		-0.54		-0.019		-0.004
		(2.379)		(0.033)		(0.004)
Natural Gas Price	0.94	-0.44	0.009	-0.009	0.002	-0.001
	(1.002)	(0.373)	(0.009)	(0.005)	(0.001)	(0.001)
Electricity Price	0.39	0.24	0.004	-0.003	0.003*	0.002
	(0.667)	(0.558)	(0.010)	(0.010)	(0.001)	(0.001)
Corn Price	-3.16	0.02	0.027	-0.022	0.009	-0.001
	(3.244)	(1.503)	(0.050)	(0.021)	(0.006)	(0.003)
Soy price	-2.70	0.13	-0.015	0.007	-0.002	0.000
	(1.417)	(0.380)	(0.025)	(0.006)	(0.003)	(0.001)
Corn Intensity	-6.27	-11.73	0.152	0.204	0.033	0.035
	(16.945)	(16.029)	(0.245)	(0.245)	(0.032)	(0.032)
Spatial Lag of Corn Intensity	-5.54	-11.48	-0.265	-0.634	0.023	-0.016
	(22.536)	(18.690)	(0.348)	(0.332)	(0.047)	(0.045)
Soy Intensity	-15.91	-17.07	-0.171	-0.128	-0.034	-0.026
	(11.944)	(10.947)	(0.174)	(0.170)	(0.024)	(0.024)
Cow Density	7.61	10.78	-0.175	-0.057	0.294***	0.302***
	(26.297)	(25.524)	(0.409)	(0.404)	(0.077)	(0.077)
Year (trend)		0.75**		0.002		0.001
		(0.250)		(0.003)		(0.000)
Constant	NO	NO	YES	YES	YES	YES
County Fixed Effects	YES	YES	YES	YES	YES	YES
Year Effects	YES	NO	YES	NO	YES	NO
Number of Observations	3,687	3,687	3,687	3,687	28,769	28,769
Number of Counties	120	120	120	120	855	855
Pseudo-R^2 or R^2	0.524	0.505	0.109	0.0989	0.0223	0.0213

Notes: Standard errors in parentheses. Significance codes: *** p<0.001, ** p<0.01, * p<0.05

Table 3: Specification tests for fixed effects logit model of ethanol plant entry

	Fixed Effects Logit		Linear Probability Models			
			FE Logit sample		Full sample	
	A	B	C	D	E	F
Hausman test of random effects vs. fixed effects (H_0: random effects preferred)						
p-value ($Pr > Chi^2$)	0.000***	0.000***	0.000***	0.000***	0.000***	0.000***
Durbin-Wu-Hausman test of endogeneity of corn intensity (H_0: corn intensity not endogenous)						
p-value ($Pr > F$)	0.687	0.883	0.589	0.882	0.933	0.768
County Fixed Effects	YES	YES	YES	YES	YES	YES
Year Effects	YES	NO	YES	NO	YES	NO
Number of Observations	3,687	3,687	3,687	3,687	28,769	28,769
Number of Counties	120	120	120	120	855	855

Significance codes: *** p<0.001, ** p<0.01, * p<0.05

Table 4: Summary statistics for the number of existing plants by ownership type

Plant Owner Type	Number of Existing Plants		
	Mean	Std. Dev.	Max
Existing Plants			
All Existing Plants	0.040	0.203	2
Singlets	0.016	0.125	1
Ethanol-Focused Firms	0.011	0.104	1
Conglomerates	0.012	0.109	1
Non-Singlets	0.023	0.150	2
Spatial Lag of Existing Plants			
All Existing Plants	0.395	0.819	8
Singlets	0.144	0.420	5
Ethanol-Focused Firms	0.102	0.383	6
Conglomerates	0.125	0.384	3
Non-Singlets	0.227	0.571	6

Table 5: Results from fixed effects logit model with number of existing plants by ownership type

Dependent variable is probability of ethanol plant entry				
	A	X	Y	Z
Existing Plants				
All	-13.79***			
	(1.72)			
Singlets		-16.05***		-13.62***
		(2.71)		(1.96)
Ethanol-Focused Firm				-10.53***
				(2.01)
Conglomerates			-14.53***	-1.96*
			(2.01)	(0.86)
Non-Conglomerates			-12.25***	
			(1.65)	
Non-Singlets		-11.25***		
		(1.69)		
Spatial Lag of Existing Plants				
All	0.57			
	(0.42)			
Singlets		0.62		-0.00
		(0.62)		(0.52)
Ethanol-Focused Firm				1.00
				(1.03)
Conglomerates			9.54***	0.64
			(1.59)	(0.92)
Non-Conglomerates			12.77***	
			(1.76)	
Non-Singlets		9.38***		
		(1.53)		
Policy Variables from Specification A	YES	YES	YES	YES
Economic Variables from Specification A	YES	YES	YES	YES
Year Effects	YES	YES	YES	YES
County Fixed Effects	YES	YES	YES	YES
Number of Observations	3,687	3,669	3,669	3,669
Number of Counties	120	119	119	119
Pseudo-R^2	0.524	0.514	0.519	0.491

Notes: Standard errors in parentheses. Significance codes: *** $p<0.001$, ** $p<0.01$, * $p<0.05$

Appendix A: Supplementary Tables

Table A.1: Robustness of fixed effects logit model to specification of corn and soy intensity and price

	Dependent variable is probability of ethanol plant entry							
	G	H	I	J	K	L	M	N
Existing Plants	-13.98***	-13.31***	-13.70***	-13.21***	-13.81***	-13.17***	-13.45***	-13.26***
	(1.798)	(1.581)	(1.701)	(1.565)	(1.749)	(1.561)	(1.643)	(1.560)
Spatial Lag of Existing Plants	0.55	0.43	0.60	0.48	0.56	0.47	0.43	0.48
	(0.427)	(0.379)	(0.417)	(0.370)	(0.427)	(0.370)	(0.418)	(0.370)
MTBE Ban	-1.02	-0.95	-1.00	-0.96	-1.06	-0.92	-0.96	-1.03
	(0.930)	(0.806)	(0.927)	(0.808)	(0.927)	(0.808)	(0.924)	(0.784)
Tax Credit	0.19	1.20*	0.06	1.06	0.10	1.03	0.45	1.12*
	(0.720)	(0.586)	(0.688)	(0.567)	(0.704)	(0.568)	(0.644)	(0.548)
RFS1		0.84		0.58		0.72		0.38
		(1.417)		(1.428)		(1.409)		(1.208)
RFS2		0.72		0.19		0.46		-0.20
		(3.003)		(3.092)		(3.022)		(2.619)
Corn Price	0.02	-9.10	0.33	-2.88	0.31	-8.01	0.53	-3.18
0.63								
	(5.201)	(0.727)	(3.276)	(1.506)	(5.064)	(0.726)	(3.279)	(0.667)
Soy Price			-2.56	0.10				
			(1.411)	(0.380)				
Ratio of Corn to Soy Price	43.94	-3.59			37.71	-3.23		
	(32.010)	(7.048)			(30.863)	(7.032)		
Corn Intensity	-4.16	-11.94	7.80	3.18	9.46	2.82	6.02	4.33
	(16.761)	(16.005)	(14.556)	(13.489)	(14.741)	(13.492)	(13.375)	(12.874)
Spatial Lag of Corn Intensity	-9.50	-12.15	-6.36	-12.14	-9.66	-12.73	-4.43	-12.68
	(22.878)	(18.733)	(22.630)	(18.671)	(22.946)	(18.714)	(22.396)	(18.468)
Soy Intensity	-16.06	-17.06						
	(11.867)	(10.922)						
Ratio of Corn to Soy Intensity			0.01	0.02	0.00	0.02		
			(0.097)	(0.090)	(0.098)	(0.089)		
Time Specification	Year Effect	Trend	Year Effect	Trend	Year Effect	Trend	Year Effect	Trend
County Fixed Effects	YES	YES	YES	YES	YES	YES	YES	YES
Number of Observations	3,687	3,687	3,687	3,687	3,687	3,687	3,736	3,736
Number of Counties	120	120	120	120	120	120	121	121
Pseudo-R^2	0.522	0.505	0.522	0.502	0.520	0.502	0.519	0.503

Notes: Standard errors in parentheses. Significance codes: *** p<0.001, ** p<0.01, * p<0.05

Table A.2: Robustness of fixed effects logit model to the specification of ethanol and gasoline prices

	Dependent variable is probability of ethanol plant entry				
	O	P	Q	R	S
Existing Plants	-13.70***	-13.08***	-13.53***	-13.48***	-13.47***
	(1.717)	(1.508)	(1.705)	(1.612)	(1.620)
Spatial Lag of Existing Plants	0.57	0.57	0.55	0.42	0.45
	(0.439)	(0.358)	(0.443)	(0.383)	(0.381)
MTBE Ban	-0.94	-1.45	-0.84	-0.95	-0.94
	(0.957)	(0.840)	(0.976)	(0.804)	(0.812)
Tax Credit	-0.02	1.60**	0.00	1.18*	1.27*
	(0.719)	(0.614)	(0.722)	(0.589)	(0.599)
RFS1		2.41*		1.11	2.18*
		(1.026)		(1.089)	(1.039)
RFS2		4.03**		1.37	4.39**
		(1.562)		(1.818)	(1.645)
Gasoline Price			-23.71	4.82**	
			(25.361)	(1.612)	
Ethanol Price					3.29**
					(1.274)
Ratio of Ethanol to Gasoline Price	-47.42**	0.11	-73.36*	0.72	-1.74
	(15.537)	(1.874)	(32.735)	(1.993)	(2.170)
Time Specification	Year Effect	Trend	Year Effect	Trend	Trend
County Fixed Effects	YES	YES	YES	YES	YES
Number of Observations	3,687	3,687	3,687	3,687	3,687
Number of Counties	120	120	120	120	120
Pseudo-R^2	0.530	0.494	0.531	0.505	0.502

Notes: Standard errors in parentheses. Significance codes: *** $p<0.001$, ** $p<0.01$, * $p<0.05$

Table A.3: Robustness of fixed effects logit model to the specification of tax credit and existing plant

	Dependent variable is probability of ethanol plant entry			
	Tax Credit Specification		Existing Plant Specification	
	T	U	V	W
Existing Plants	-13.72***	-13.61***		
	(1.716)	(1.617)		
Existing Ethanol Capacity [gal per acre]			-4.53***	-4.63***
			(0.528)	(0.533)
Existing Biodiesel Plant	-0.49	-0.16		
	(1.704)	(1.454)		
Existing Biodiesel Capacity [gal per acre]			0.01	0.01
			(0.037)	(0.035)
MTBE Ban	-0.89	-1.08	-0.41	-0.75
	(0.926)	(0.821)	(0.837)	(0.710)
Tax Credit (indicator)			1.04	1.45*
			(0.615)	(0.564)
Lifetime Tax Credit Benefit ($100,000)	0.04	0.10*		
	(0.047)	(0.040)		
RFS1		0.85		0.16
		(1.438)		(1.265)
RFS2		0.54		0.11
		(3.078)		(2.880)
Time Specification	Year Effect	Trend	Year Effect	Trend
County Fixed Effects	YES	YES	YES	YES
Number of Observations	3,687	3,687	3,687	3,687
Number of Counties	120	120	120	120
Pseudo-R^2	0.525	0.507	0.480	0.464

Notes: Standard errors in parentheses. Significance codes: *** p<0.001, ** p<0.01, * p<0.05