Ethanol Plant Investment and Government Policy: A Dynamic Structural Econometric Model*

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

December 30, 2018

Abstract

Ethanol has attracted considerable policy 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 effects of government policy on the decisions of ethanol-producing firms to invest in building new ethanol plants in the Midwestern United States during the second US ethanol boom. To do so, we develop and estimate a dynamic structural econometric model of the ethanol plant investment timing game. According to our results, government policies, the intensity of corn production, and private information shocks all have significant effects on ethanol investment payoffs and decisions. We use the estimated structural parameters to simulate counterfactual policy scenarios to disentangle the impacts of state and national policies on the timing and location of investment in the industry. We find that, of the policies analyzed, the policies that led to most of the ethanol plant investment during this time period were the ban on the use of the oxygenate MTBE as a gasoline additive, and the 2007 Renewable Fuel Standard (RFS2).

Keywords: ethanol, investment timing game, structural model
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 policy attention as an environmentally-friendly alternative to imported oil; as a substitute, additive, oxygenate, and/or octane booster for gasoline (Irwin and Good, 2017); and as a way to boost farm profits and improve rural livelihoods. In the United States, government policies that actively promote ethanol production have coincided with a boom in the construction of corn-ethanol plants, known as the second US ethanol boom, that began in the mid-1990s and hit full-stride by the early 2000s.[1]

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.

In this paper, we analyze the investment decisions of ethanol-producing firms in the Midwestern United States during the second US ethanol boom. The decision to invest in building an ethanol plant is a dynamic decision that may be affected by economic factors and government policies. Because the payoff from investing in building a new ethanol plant depend on market conditions such as the feedstock price that vary stochastically over time, a potential entrant that hopes to make a dynamically optimal decision would need to account for the option value to waiting before making this irreversible investment (Dixit and Pindyck, 1994). For example, commodity markets occasionally exhibit broadly based massive booms and busts; at the core of these cycles is a set of contemporaneous supply and demand surprises

[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).

that coincide with low inventories and that are magnified by macroeconomic shocks and policy responses (Carter, Rausser and Smith, 2011). Market volatility can induce periods of boom and bust in the ethanol industry, causing episodes of bankruptcy and reduced capital investment (Hochman, Sexton and Zilberman, 2008).

The dynamic decision-making problem faced by a potential ethanol investor is even more complicated when the investment payoff is affected not only by market conditions and government policies, but also by the existence of nearby plants. Due to potential competition effects and agglomeration effects (Lin Lawell, 2017; Thome and Lin Lawell, 2018; Yi and Lin Lawell, 2018a; Yi and Lin Lawell, 2018b), the presence of existing ethanol plants may affect the payoff from investing in an ethanol plant. Because the investment decisions of other potential investors affect the future values of state variables and the future payoff from investing in a new ethanol plant, potential ethanol investors must anticipate the investment strategies of other potential investors in order to make a dynamically optimal decision. As a consequence, a potential ethanol investor's investment decision depends on its conjecture about competitors' behavior. Uncertainty over whether a plant might be constructed and start production nearby is another reason there is an option value to waiting before investing that makes the decision dynamic rather than static (Dixit and Pindyck, 1994).

In this paper, we estimate a structural econometric model of the ethanol plant investment timing game. We use the estimated parameters from the structural model to run counterfactual simulations to explore the policy factors driving industry growth and location.

There are several advantages to using a structural approach to analyzing the decision to invest in building a new ethanol plant. First, our structural model explicitly models the dynamic investment decision, including the continuation value to waiting. A potential entrant invests if the payoff from investment exceeds the continuation value from waiting.

A second advantage of our structural model is that we are able to estimate the effect of each state variable on the expected payoff from investing in an ethanol plant. While the parameters in reduced-form models are confounded by continuation values, we model the structural relationship between the continuation value from waiting and the payoff from investment, which enables us to estimate parameters in the payoff from investing in building an new ethanol plant.

A third advantage of a structural model is that the parameter estimates from the structural model can be used to simulate counterfactual scenarios. We use the estimated parameters from the structural model to run counterfactual simulations to explore the effects of alternative policies on ethanol investment.

According to our results, government policies, particularly the ban on the use of the oxygenate MTBE as a gasoline additive, and the 2007 Renewable Fuel Standard (RFS2), have significant effects on ethanol investment payoffs and decisions. The intensity of corn production and private information shocks have significant effects on ethanol investment payoffs and decisions as well. We use the estimated structural parameters to simulate counterfactual policy scenarios to disentangle the impacts of state and national policies on the timing and location of investment in the industry. We find that, of the policies analyzed, the MTBE ban and the RFS2 led to most of the ethanol plant investment during this time period.

The balance of our paper proceeds as follows. In Section 2, we review the relevant literature. We present our structural econometric model in Section 3. We describe our data in Section 4. We present our results in Section 5. We run counterfactual simulations in Section 6. Section 7 concludes.

2 Literature Review

2.1 Ethanol investment and location decisions

The first branch of literature on which we build is that on models of business investment and location decisions. For excellent reviews of this literature, see Bartik (1985) and Goetz (1997). In empirical models of business investment and location decisions, investment in businesses, particularly manufacturing, is often modeled as a function of output market prices and access, input costs and access, and the policy environment.

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. 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. 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. Thome and Lin Lawell (2018) analyze the effects of local competition and agglomeration on ethanol plant entry decisions.

The location determinants identified in these studies provide a starting point for our analysis as far as identifying potentially important state variables to include in our structural model. 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.

We improve upon these previous models by estimating a dynamic structural econometric model with panel data, by analyzing investment timing decisions, by directly estimating the effect of covariates on the payoff to investment, and by using the estimated structural parameters to simulate entry decisions and welfare under various counterfactual policy scenarios.

2.2 Ethanol investment and the effects of 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).

As for studies of the Renewable Fuel Standard, a number of authors have studied renewable fuel mandates and their effects 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). Lade and Lin Lawell (2018) develop a theory model of renewable fuel mandates and apply it to the Renewable Fuel Standard. Lade, Lin Lawell and Smith (2018b) develop a dynamic model of Renewable Fuel Standard compliance. Korting and Just (2017) develop a model of the Renewable Fuel Standard that accounts for nested mandates and explores four fundamental channels of mandate compliance. Lade, Lin Lawell and Smith (2018a) draw lessons from the Renewable Fuel Standard for the design of climate policy. Stock (2015, 2018) considers and examines possible regulatory and legislative reforms to the Renewable Fuel Standard.

We build upon these previous models by estimating investment strategies econometrically, by directly estimating the effect of covariates on the payoff to investment, and by using the estimated structural parameters to simulate entry decisions and welfare under various counterfactual policy scenarios.

2.3 Dynamic structural econometric models

A third branch of literature upon which we build is that on dynamic structural econometric modeling. Rust's (1987, 1988) seminal papers develop a dynamic structural econometric model using nested fixed point maximum likelihood estimation. Hotz et al. (1994) develop a conditional choice simulation estimator for dynamic models of discrete choice. Dynamic structural econometric models have been adapted for many applications, including bus engine replacement (Rust, 1987), nuclear power plant shutdown (Rothwell and Rust, 1997), water management (Timmins, 2002), agriculture (Scott, 2013), air conditioner purchases

(Rapson, 2014), wind turbine shutdowns and upgrades (Cook and Lin Lawell, 2018), copper mining decisions (Aguirregabiria and Luengo, 2016), crop disease control (Carroll et al., 2018b), vehicle scrappage programs (Li and Wei, 2013), the adoption of rooftop solar photovoltaics (Feger et al., 2017; Langer and Lemoine, 2018), supply chain externalities (Carroll et al., 2018a), organ transplant decisions (Agarwal et al., 2018), pesticide spraying decisions (Sambucci, Lin Lawell and Lybbert, 2018), vehicle ownership and usage (Gillingham et al., 2016), insecticide treated nets (Mahajan and Tarozzi, 2011), and agricultural productivity (Carroll et al., forthcoming).

We build in particular on the literature on structural econometric models of dynamic games. Some examples are Aguirregabiria and Mira (2007), who develop structural econometric methods for sequential estimation of dynamic games; and Bajari, Benkard and Levin (2007), who develop a structural econometric model of a dynamic game with continuous control variables. Structural econometric models of dynamic games have been applied to environmental regulation (Ryan, 2012), fisheries (Huang and Smith, 2014), market-based emissions regulation (Fowlie, Reguant and Ryan, 2016), utility regulation (Lim and Yurukoglu, 2018), Chinese shipbuilding (Kalouptsidi, 2018), the world petroleum market (Kheiravar, Lin Lawell and Jaffe, 2018), the global market for solar panels (Gerarden, 2018), subsidies (Yi, Lin Lawell and Thome, 2018), industrial policy (Barwick, Kalouptsidi and Zahur, 2018), coal procurement (Jha, 2018), migration decisions (Rojas Valdés, Lin Lawell and Taylor, 2018a; Rojas Valdés, Lin Lawell and Taylor, 2018b), the digitization of consumer goods (Leyden, 2018), and climate change policy (Zakerinia and Lin Lawell, 2018).

In this paper, we apply a structural econometric model of a dynamic game developed by Pakes, Ostrovsky and Berry (2007). This model has been applied to analyze the multi-stage investment timing game in offshore petroleum production (Lin, 2013), and to peer effects in health promotion programs in developing countries (Ma, Lin Lawell and Rozelle, 2018).

3 Structural Econometric Model

We model the dynamic and strategic decision faced by a potential investor (or entrant)[2] i of whether to invest in building an ethanol plant in county k in year t. Investment in an ethanol plant is irreversible and, in each year t, all investment decisions are made simultaneously. I_{ikt} is an indicator of whether potential investor i invests in building a new ethanol plant in county k in year t.

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

The publicly observable state of county k in year t is given by $\Omega_{kt} = (N_{kt}, G_{kt}, X_{kt})$, a vector of discrete and finite-valued state variables that are observed by all the potential investors in county k as well as by the econometrician, where N_{kt} is a dummy variable for whether there is an existing plant in the county; G_{kt} describes the policy environment; and X_{kt} are economic factors. The state variables N_{kt}, G_{kt}, and X_{kt} evolve according to a first-order controlled Markov process and summarize the direct effect of the past on the current environment.

The payoff $\pi(\cdot)$ from investing in an ethanol plant in year t, which represents the present discounted value of the entire stream of net benefits from investing in an ethanol plant, depends on the state variables Ω_{kt} at time t. As a consequence, the decision of a potential investor i of whether to invest in building an ethanol plant in each county k in year t depends on the state variables Ω_{kt} as well.

The value function for a potential investor i in county k in period t is the expected present discounted value of the entire stream of net benefits to the potential investor from following the dynamically optimal investment policy, and can be written as:

$$V(N_{kt}, G_{kt}, X_{kt}) = max\{\pi(N_{kt}, G_{kt}, X_{kt}), \beta V^c(N_{kt}, G_{kt}, X_{kt})\}, \tag{1}$$

where β is the discount factor. The continuation value $V^c(\cdot)$ 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:

$$V^c(N_{kt}, G_{kt}, X_{kt}) = E[V(N_{kt+1}, G_{kt+1}, X_{kt+1})|N_{kt}, G_{kt}, X_{kt}, I_{ikt} = 0], \tag{2}$$

where expectations are taken over the values of the next period's state variables conditional on the value of the current period's state variables and conditional on not investing in the current year.

In a static model of investment, the statically optimal investment rule is to invest if the payoff $\pi(\cdot)$ from investing is greater than 0. When investments are irreversible and there is uncertainty over the future payoff from investment, however, the statically optimal investment rule is not dynamically optimal. In particular, if the state variables $\Omega_{kt} = (N_{kt}, G_{kt}, X_{kt})$ evolve stochastically over time, then it is possible that the state variables may take on values in the future that yield a payoff $\pi(\cdot)$ that is high enough that the potential investor would do better in expected present discounted value to wait rather than make the investment now, even if the payoff $\pi(\cdot)$ now is positive. A potential investor holds an option to invest, which it loses when the irreversible investment is made. This lost option value is an opportunity cost that must be included as part of the cost of investment. A potential investor who hopes to make a dynamically optimal decision would therefore need to account for the option value to waiting before making this irreversible investment (Dixit and Pindyck, 1994).

The dynamically optimal investment policy is for the potential entrant to build an ethanol plant in year t if and only if the payoff $\pi(N_{kt}, G_{kt}, X_{kt})$ from investing exceeds β times the continuation value to waiting, $V^c(\cdot)$:

$$I_{ikt} = \begin{cases} 1 & \text{if } \pi(N_{kt}, G_{kt}, X_{kt}) > \beta V^c(N_{kt}, G_{kt}, X_{kt}) \\ 0 & \text{if } \pi(N_{kt}, G_{kt}, X_{kt}) \leq \beta V^c(N_{kt}, G_{kt}, X_{kt}). \end{cases} \tag{3}$$

Because the continuation value from waiting is positive, the dynamically optimal investment rule has a higher threshold for the payoff from investment to exceed before an investment is made compared to the static investment rule, whose threshold is 0. Because of the uncertainty in state variables measuring economic factors and government policy, there is an option value to waiting, which means that potential entrants are more likely to delay their investments. This uncertainty combined with the irreversible nature of ethanol plant investment make a dynamic model more appropriate than a static model to model ethanol plant investment. Thus, our structural model, which is dynamic, is more appropriate than a reduced-form model, which does not explicitly model the continuation values.

The dynamic decision-making problem faced by a potential investor is even more complicated when the investment payoff is affected not only by market conditions and government policies, but also by the existence of nearby plants. Due to potential competition effects and agglomeration effects (Lin Lawell, 2017; Thome and Lin Lawell, 2018; Yi and Lin Lawell, 2018a; Yi and Lin Lawell, 2018b), the presence of existing ethanol plants may affect the payoff from investing in an ethanol plant. As a consequence, a potential investor's investment decision depends on its conjecture about competitors' behavior.

In particular, potential entrants may condition their investment decisions on both whether there is an existing ethanol plant in the county N_{kt} and their expectations on what future values of N_{kt} may be. Future values of N_{kt} may be different from current values if other potential entrants enter in a given year. In our model, potential entrants base their decisions in part on expectations of the future, including their expectations of whether a plant will be built in county k by the next year, which depend on what they expect other potential entrants to do in a given period. By structurally capturing a potential entrant's beliefs about other potential entrants, we are unable to structurally model the effect of other potential entrants on a potential entrant's payoffs.

The state variables in G_{kt} describe the policy environment faced by the corn-ethanol industry. State and federal policies can affect the expected payoff from investing in building a new ethanol plant 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 tax credits for ethanol producers.

The state variables in X_{kt} include economic factors that affect the payoffs from investing in building an ethanol plant. 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).[3] The gasoline price could have a positive or negative impact on investment depending on whether ethanol is viewed as an energy substitute for gasoline or as a gasoline additive (oxygenate and/or octane booster), respectively.

The vector X_{kt} of economic factors also includes state variables 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). 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 include a metro area indicator in order to capture proximity to market and transportation costs.[4]

The vector X_{kt} of economic factors also includes soy price and 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.

In addition to the observable state variables $\Omega_{kt} = (N_{kt}, G_{kt}, X_{kt})$, the decision of a potential investor i of whether to invest in building an ethanol plant in each county k in year t also depends on a shock ε_{ikt}, which is private information to the potential investor and unobserved by either other potential investors or by the econometrician. Such private

[3]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 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).

[4]The modeling of transportation infrastructure investment decisions, which affect transportation costs and may be endogenous at the county level, and which has been studied elsewhere (Fatal et al., 2012), is beyond the scope of this paper.

information may include, for example, a shock to the cost of building an ethanol plant. We assume the error term is independently and identically distributed exponentially with mean σ, which is among the parameters to be estimated.

The equilibrium concept used in the model is that of a Markov perfect equilibrium. Each potential investor is assumed to play a Markov "state-space" strategy: the past influences current play only through its effect on the state variables. A potential investor's dynamically optimal investment policy is then the Markov strategy that it plays in the Markov perfect equilibrium, which is a profile of Markov strategies that yields a Nash equilibrium in every proper subgame (Fudenberg and Tirole, 1998).

While each potential investor's time-t investment decision depends on both the publicly available endogenous and exogenous state variables Ω_{kt} as well as the potential investor's own private information ε_{ikt}, its perception of other potential investors' time-t investment decisions depend only on the publicly observable state variables Ω_{kt}. This is because, owing to the above assumptions on the observable state variables and on the unobservable shocks, potential investors can take expectations over their competitors' private information.[5] In equilibrium, potential investors' perceptions of their competitors' investment probabilities should be consistent with those that are actually realized (Starr and Ho, 1969).

The model has at least one Markov perfect equilibrium, and each equilibrium generates a finite state Markov chain in Ω_{kt} tuples (Pakes, Ostrovsky and Berry, 2007).[6] Although model assumptions do not guarantee a unique equilibrium, they do insure that there is only one set of equilibrium policies that is consistent with the data generating process. It is thus possible to use the data itself to pick out the equilibrium that is played. For large enough samples, the data will pick out the correct equilibrium and the estimators for the parameters in the model will be consistent (Pakes, Ostrovsky and Berry, 2007).[7]

The payoff $\pi(N_{kt}, G_{kt}, X_{kt}, \varepsilon_{ikt}; \theta)$ from investing in an ethanol plant in county k in year t can be separated into a deterministic component and a stochastic component as follows:

$$\pi(N_{kt}, G_{kt}, X_{kt}, \varepsilon_{ikt}; \theta) = \pi_0(N_{kt}, G_{kt}, X_{kt}; \theta) + \varepsilon_{ikt}, \tag{4}$$

where the deterministic component $\pi_0(\cdot)$ is linear in the state variables:

$$\pi_0(N_{kt}, G_{kt}, X_{kt}; \theta) = N'_{kt}\gamma_N + G'_{kt}\gamma_G + X'_{kt}\gamma_X, \tag{5}$$

[5]While each potential investor plays a pure strategy, from the point of view of their competitors, they appear to play mixed strategies. Thus, as with Harsanyi's (1973) purification theorem, a mixed distribution over actions is the result of unobserved payoff perturbations that sometimes lead potential investors to have a strict preference for one action, and sometimes a strict preference for another.

[6]A Markov chain is a Markov process on a finite state space (Stokey, Lucas and Prescott, 1989).

[7]This assumes that the same equilibrium is played in each market. If a mixed strategy equilibrium is played, then it is assumed that the same mixed strategy equilibrium is played in each market.

and where $\theta = (\gamma_N, \gamma_G, \gamma_X, \sigma)$ denotes the parameters to be estimated. The coefficients γ_N, γ_G, and γ_X measure the effects of the state variables N_{kt}, G_{kt}, and X_{kt}, respectively, on the payoff to investing in building a new ethanol plant.

The value function for a potential entrant i in county k in period t can be written as:

$$V(N_{kt}, G_{kt}, X_{kt}, \varepsilon_{ikt}; \theta) = max\{\pi(N_{kt}, G_{kt}, X_{kt}\varepsilon_{ikt}; \theta), \beta V^c(N_{kt}, G_{kt}, X_{kt}; \theta)\}. \quad (6)$$

The potential entrant will invest if and only if the payoff from investing exceeds β times the continuation value to waiting. The continuation value, $V^c(\cdot)$, 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:

$$V^c(N_{kt}, G_{kt}, X_{kt}; \theta) = E[V(N_{kt+1}, G_{kt+1}, X_{kt+1}, \varepsilon_{ikt+1}; \theta)|N_{kt}, G_{kt}, X_{kt}, I_{ikt} = 0]. \quad (7)$$

Let $g(N_{kt}, G_{kt}, X_{kt}; \theta)$ denote the probability of investing in an ethanol plant at time t, conditional on the publicly available information $\Omega_{kt} = (N_{kt}, G_{kt}, X_{kt})$ at time t, but not on the private information ε_{ikt}. The investment probability function $g(N_{kt}, G_{kt}, X_{kt}; \theta)$ represents a potential investor's perceptions of the probability that a competitor who has not yet invested will decide to invest at time t.

Using the exponential distribution for ε_{ikt} the continuation value $V^c(\cdot)$ reduces to:

$$\begin{aligned} V^c(N_{kt}, G_{kt}, X_{kt}; \theta) = E[&\beta V(N_{kt+1}, G_{kt+1}, X_{kt+1}, \varepsilon_{ikt+1}; \theta) \\ &+ \sigma g(N_{kt+1}, G_{kt+1}, X_{kt+1}; \theta)|N_{kt}, G_{kt}, X_{kt}, I_{ikt} = 0], \end{aligned} \quad (8)$$

and the investment probability $g(\cdot)$ reduces to:

$$g(N_{kt}, G_{kt}, X_{kt}; \theta) = exp\left(-\frac{\beta V^c(N_{kt}, G_{kt}, X_{kt}; \theta) - \pi_0(N_{kt}, G_{kt}, X_{kt}; \theta)}{\sigma}\right), \quad (9)$$

as shown by Lin (2013).

We employ a two-step semi-parametric estimation procedure following Pakes, Ostrovsky and Berry (2007) and Lin (2013). In the first step, the continuation value is estimated non-parametrically and this estimate is used to compute the predicted probabilities of investment. In the second step, the parameters $\theta = (\gamma_N, \gamma_G, \gamma_X, \sigma)$ are estimated by matching the predicted probabilities with the actual probabilities in the data using generalized method of moments (GMM).

For the first step in the estimation we construct a transition matrix *M*, which describes the evolution of the state variables N_{kt}, G_{kt}, and X_{kt} over time, conditional on not investing. The transition matrix M gives, for each combination of state variables in year t, the probability of transitioning to each combination of state variables in year *t+1* conditional on not investing in year t. The element in each row r, column c is represented by: $M_{rc} = Pr(\Omega_{k,t+1} = c|\Omega_{kt} = r, I_{ikt} = 0)$. We estimate M non-parametrically using empirical averages. We therefore assume rational expectations on the part of potential ethanol investors, namely that their expectations about the evolution of state variables over the time period of our data set were consistent with the actual evolution realized.

Let $\bar{g}$ be the vectorized investment policy function, which is a vector whose length is the number of combinations of state variables and whose value at each component is the investment policy function $g(\cdot)$ evaluated at a particular combination of state variables. $\bar{g}$ gives the probability of investment in a new ethanol plant for every observed state of the world $\Omega_{kt} = (N_{kt}, G_{kt}, X_{kt})$. We estimate $\bar{g}$ using empirical averages:

$$\bar{g}(N_{kt}, G_{kt}, X_{kt}) = Pr(I_{ikt} = 1|N_{kt}, G_{kt}, X_{kt}). \tag{10}$$

From Equation (8), the vectorized continuation value $\bar{V}^c$, which is a vector whose length is the number of combination of state variables and whose value at each component is the continuation value V^c evaluated at a particular combination of state variables, can be specified in vector form as $\bar{V}^c = M(\beta\bar{V}^c + \sigma\bar{g})$, where M is the empirical transition matrix, β is the discount rate, and $\bar{g}$ is the vector of empirical investment probabilities. Because this is an infinite horizon problem, we estimate the continuation value by solving for the fixed point $\hat{V}^c$, which, from Blackwell's Theorem, is unique. We then use the estimate $\hat{V}^c$ to form the predicted probability of investment in an ethanol plant, which from Equation (9) can be specified in vector form as:

$$\hat{g}(N_{kt}, G_{kt}, X_{kt}; \theta) = exp\left(-\frac{\beta\hat{V}^c - N'_{kt}\gamma_N - G'_{kt}\gamma_G - X'_{kt}\gamma_X}{\sigma}\right). \tag{11}$$

In the second step of the estimation procedure, we estimate the parameters $\theta = (\gamma_N, \gamma_G, \gamma_X, \sigma)$ by finding the parameters that best match the investment probability predicted by our model with the respective empirical investment probabilities in the data using GMM. We use the following moment function:

$$\psi = (\hat{g}(N_{kt}, G_{kt}, X_{kt}; \theta) - \bar{g}(N_{kt}, G_{kt}, X_{kt}))n(N_{kt}, G_{kt}, X_{kt}|I_{ikt-1} = 0), \tag{12}$$

where $n(N_{kt}, G_{kt}, X_{kt}|I_{ikt-1} = 0)$ counts the number of times each state $\Omega_{kt} = (N_{kt}, G_{kt}, X_{kt})$

occurs where there is a potential investor. 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_{kt}, 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. Additional moments are constructed by interacting the above moments ψ with the state variables Ω_{kt}.

The GMM estimator $\hat{\theta}$ is the solution to the problem:

$$\min_{\theta} \left(\frac{1}{obs} \sum \psi \right) W_n^{-1} \left(\frac{1}{obs} \sum \psi \right), \tag{13}$$

where obs is the number of investor-county-year observations. Because the system is exactly identified, we use an identity matrix as the weight matrix W_n.[8]

We form standard errors by a nonparametric bootstrap. We randomly draw counties from the data with replacement to generate 250 independent panels of size equal to the actual sample size. The structural econometric model is run on each of the new panels. The standard error is then formed by taking the standard deviation of the estimates from each of the random samples.

The problem of spatially correlated unobservables can be addressed by interpreting the investment payoff in the model as expected investment payoff conditional on observables, where the expectation is taken over the correlated unobservables. The model is still able to separately identify the (expected) strategic interaction from the correlated unobservables. The online Appendix of Lin's (2013) Monte Carlo experiments analyzes the effect of a state variable that is observed by firms when they make their decisions but unobservable to the econometrician (i.e., a common shock), and show that the bias introduced by spatially correlated unobservables is small. This is consistent with Pakes, Ostrovsky and Berry (2007), who find that the bias from serially correlated common shocks is small.

4 Data

4.1 Time Frame and Focus Region

We focus on investments in 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

[8]One challenge is determining whether the model has converged at a global or local minimum. We experimented with several combinations of starting values to initialize the parameters to be estimated. We found the model is robust to the starting value.

on corn-ethanol plants eliminates the need to consider feedstock choice in the model.[9] 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. For our econometric analysis, we eliminate completely non-agricultural counties within the ten states (e.g. northern Minnesota), as well as those with missing data on agricultural production.

We focus on the time period 1996 to 2008, which corresponds to the second 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.[10] 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.[11] Finally, much of the data on other variables are publicly available at an annual level.

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,

[9]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.

[10]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.

[11]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.

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 investment variable I_{ikt} is an indicator of whether potential investor i invests in 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 investors per county-year. The investment variable I_{ikt} is equal to 1 if the plant enters in a given calendar year.[12] Once a potential investor i invests, it is no longer a potential investor and therefore exits the sample.

The dummy for existing plants N_{kt} in the county is a dummy variable for whether there is an operational plant in that county on January 1 of year t, and is therefore observable to any potential investor making a decision in year t.

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 two 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.[13] 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.

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

[13]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.

For federal-level policy variables, we specify two variables to capture the effects of the Renewable Fuel Standards (RFS).[14] 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 Economic Variables

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. 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 dummy variable N_{kt} for whether there is an existing ethanol plant 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.[15] The potential DDGS market also includes hogs, but data is not available at the district level for all states. Nevertheless, 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

[14] 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.

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

at state level.[16] 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 dummy variable N_{kt} for whether there is an existing ethanol plant in the county.

4.5 State Variables

The publicly observable state of county k in year t is given by $\Omega_{kt} = (N_{kt}, G_{kt}, X_{kt})$, a vector of discrete and finite-valued state variables that are observed by all the potential investors in county k as well as by the econometrician, where N_{kt} is a dummy variable for whether there is an existing plant in the county; G_{kt} describes the policy environment; and X_{kt} are economic factors.

We discretize each of the continuous variables in our data into discrete and finite-valued state variables, as detailed in Table 1. For our base specification, we discretize the continuous variables into two bins each. In some cases we aim for equally-sized bins (*natural gas price*, *electricity price*, *gasoline price*, *ethanol price*, *corn intensity*). For other variables, owing in part to their skewed distribution, we create bins that put higher weight on the lower (*corn price*) or higher (*cow density*) part of the continuous variable. We also construct alternate bins to test the robustness of our model to different break points, including discretizing the continuous variables into three instead of two bins. Summary statistics for the discretized state variables used in our structural model are in Table 2.

Each state of the county $\Omega_{kt} = (N_{kt}, G_{kt}, X_{kt})$ is represented by a combination of discretized state variables. The number of potential states of the county is the product of the number of bins of each state variable. Dimensionality is an important consideration for the simulations we perform using the structural estimates. For example, when we simulate removing a policy, we must observe the rest of the variables describing the state of the world Ω_{kt} with and without the policy. Thus, our preferred specification has fewer bins and covariates, thus fewer potential states of the world Ω_{kt} that we must identify and observe to conduct simulations.

Because the main objectives of this paper are to learn about the effects of government policy on investment, we are most concerned with the other state variables to the extent that they can fully describe the state of input and output markets. The indicator variables

[16] 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.

we construct for output and energy input prices allow us to control for prices in the state of the world, while freeing up dimensions to focus on and identify different policies in our simulations. The variable *energy input price* is an indicator that is one when both the electricity and natural gas prices are high. The variable *output price* indicator is one when both the gasoline and ethanol price is high.

5 Results of Structural Model

The results from the structural estimation of the parameters are reported in Tables 3 and 4. Our preferred specification, which we use for the counterfactual policy simulations, is specification (i). Our preferred specification includes only *natural gas price* and not *electricity price* or the *input price indicator* because the reduced-form analysis in Thome and Lin Lawell (2018) indicates that *electricity price* does not have a significant impacts on the probability of entry. The additional specifications (ii)-(vi) in Table 3 show the robustness of the model to different price specifications. Table 4 has an alternate specification with additional covariates whose effects we cannot separately identify for the policy simulations (specification (vii)) and also shows the results with alternate bins (specifications (viii)-(ix)).

All of the policy variables have positive impacts on the payoff from investment in an ethanol plant, and two, *MTBE Ban* and *RFS2*, are significant. Because both the MTBE ban and the Renewable Fuel Standard can function as implicit blending mandates (de Gorter and Just, 2010; Anderson and Elzinga, 2014), the similar magnitude of the coefficients suggests similar implicit state blending levels. Further, the coefficient on *RFS1* is much smaller and is not statistically significant, which would suggest that the first version of the RFS was not big enough to induce investment.

On the input (cost) side, *corn intensity* has a positive impact on the payoff from investment, while *corn price* is not significant. This result is similar to the reduced-form literature on plant location, which finds that physical access to feedstock is a significant location determinant, but more aggregate feedstock prices are not important (e.g. Cotti and Skidmore, 2010).

On the revenue side, the coefficient on *output price indicator* is negative; this means when both ethanol and gasoline prices are high, there is a negative impact on the payoff from investing. In the alternate price specifications (iii) and (iv) in Table 3, we show that high ethanol and gasoline prices have negative impacts on the payoff from investment when modeled individually, though the effects are insignificant.

Specifications (v)-(vi) in Table 3 show the robustness of the model to various specifications of the input price variables, none of which have significant impacts on the payoffs from investing in an ethanol plant.

The constant and the mean of the private shock σ are both significant determinants of the payoff from investing. The estimate of σ is similar in magnitude to the coefficients on *MTBE ban* and *RFS2*, indicating that this private information shock can be as important as the policies in determining investment payoff. The constant is large and negative, indicating there are significant fixed costs to investing in an ethanol plant.

Specification (vii) in Table 4 builds on the base specification by adding the additional covariates *metro area* and *biodiesel plant*. These variables have insignificant effects on the expected payoff from investing in an ethanol plant, and their inclusion does not lead to noticeable differences in the other estimates. Consequently, we do not include these covariates in our preferred specification.

Specifications (viii) and (ix) in Table 4 show the results of structural estimation with alternate bins and more covariates than our preferred specification (specification (i)). Since dimensionality is an important consideration for the simulations we perform using the structural estimates, however, our preferred specification (i) has fewer bins and covariates, thus fewer potential states of the world Ω_{kt} that we must identify and observe to conduct simulations.

6 Counterfactual Simulations

We conduct counterfactual simulations to assess the goodness of fit our model and to analyze the effects of counterfactual government policies.

To assess the goodness of fit of our model, we conduct a replication exercise in which we use our estimated model applied to the observed exogenous state variables to simulate (or predict) the data. We call the model predicted results our Base scenario.

In addition to the simulated replication using the observed exogenous state variables (the Base scenario), we simulate investment under several counterfactual policy scenarios. These counterfactual scenarios include No RFS1, No RFS2, No Tax Credit, No MTBE Ban, and No Policy, and are summarized in Table 5.

6.1 Methodology

The methodology for the counterfactual simulations and replication exercise begins by using the observed data to construct the empirical investment probabilities $\bar{g}(\cdot)$ and a transition matrix M for each state of the world $\Omega_{kt} = (N_{kt}, G_{kt}, X_{kt})$ using empirical averages; these are identical to the vectors used in the structural estimation. Substituting in the structural model estimate of $\hat{\sigma}$ into Equation (8), we calculate the estimated continuation value for each state of the world $\Omega_{kt} = (N_{kt}, G_{kt}, X_{kt})$ by solving for a fixed point $\tilde{V}^c(\cdot)$:

$$\tilde{V}^c(N_{kt}, G_{kt}, X_{kt}, ; \hat{\theta}) = M(\beta\hat{V} + \hat{\sigma}\bar{g}(\cdot)). \tag{14}$$

This vector has one observation for each of the observed states of the world Ω_{kt}.

We next substitute the structural parameters estimates, $\hat{\theta} = (\hat{\gamma}_N, \hat{\gamma}_G, \hat{\gamma}_X, \hat{\sigma})$ from specification (i) of Table 3, and the estimated continuation value $\tilde{V}^c$ from Equation (14) into the expression for the predicted probability of investment from Equation (11) to form the estimated probability of investment $\tilde{g}(\cdot)$, where:

$$\tilde{g}(N_{kt}, G_{kt}, X_{kt}; \hat{\theta}) = exp\left(-\frac{\beta\tilde{V} - N'_{kt}\hat{\gamma}_N - G'_{kt}\hat{\gamma}_G - X'_{kt}\hat{\gamma}_X}{\hat{\sigma}}\right). \tag{15}$$

Each observed state of the world $\Omega_{kt} = (N_{kt}, G_{kt}, X_{kt})$ observed in that data is associated with an estimated probability of entry in the vector $\tilde{g}(\cdot)$.

We begin simulating investment at *t=1*, which corresponds to 1996, our first year of data. In the first step, for each county k, we evaluate the estimated probability of entry $\tilde{g}(N_{kt}, G_{kt}, X_{kt}; \hat{\theta})$ for the state of the world $\Omega_{kt} = (N_{kt}, G_{kt}, X_{kt})$ for county k at time t. For the replication, we use the observed exogenous state variables to define Ω_{kt}.

In the second step of the simulation, we take a random draw, d, from a uniform distribution for each potential entrant i so that $d_{ikt} \sim U(0, 1)$. The entry rule is that a potential investor enters if the random draw is less than the estimated probability of investment in the current state of the world:

$$I^d_{ikt} = \begin{cases} 1, & \text{if } d_{ikt} < \tilde{g}(N_{kt}, G_{kt}, X_{kt}; \hat{\theta}) \\ 0, & \text{if } d_{ikt} \geq \tilde{g}(N_{kt}, G_{kt}, X_{kt}; \hat{\theta}). \end{cases} \tag{16}$$

Once a potential investor i makes an investment ($I_{ikt} = 1$), that investor exits the sample.

In the third simulation step, we update N_{kt} for year t+1 to account for any investments made in each county k in year t. We use the observed data for the exogenous variables G_{kt} and X_{kt}. This implies that the evolution of these state variables is not dependent on the number of ethanol plants in a county, N_{kt}. Many of the variables, including all of the policies, are measured at aggregate levels that include many counties k and potential investors i. Thus while they move together, we would not expect a rare county-level event to affect their trajectory. Other than the number of competing plants, the only other variable measured at the county level is corn intensity.

We repeat steps 1-3 for each year through 2008 (the 13th year), updating N_{kt} for each period.

In the final step, after simulating entry for the period 1996-2008, we record the total number of entrants, E, and the number of entrants in each year t, E_t. We also calculate the total welfare of all entrants, W, and the welfare of each entrant, w_e, which is specified as:

$$w_e = E[\pi(N_{kt}, G_{kt}, X_{kt}, \varepsilon_{ikt}; \hat{\theta})|N_{kt}, G_{kt}, X_{kt}] = N'_{kt}\hat{\gamma}_N + G'_{kt}\hat{\gamma}_G + X'_{kt}\hat{\gamma}_X + \hat{\sigma}. \tag{17}$$

The expression for entrant welfare w_e incorporates both the deterministic part of the payoff from investing, $\pi_0(N_{kt}, G_{kt}, X_{kt}; \hat{\theta}) = N'_{kt}\hat{\gamma}_N + G'_{kt}\hat{\gamma}_G + X'_{kt}\hat{\gamma}_X$, as well the mean of the private shock $E[\varepsilon_{ikt}; \hat{\theta})] = \hat{\sigma}$.

The mean welfare per entrant $\bar{w}_e$ is taken over all entrants e in all years for a given simulation, and is given by:

$$\bar{w}_e = \frac{\sum_{e=1}^{E} w_e}{E}. \tag{18}$$

The standard deviation of the welfare per entrant, s_e, over all years of the simulation is:

$$s_w = \sqrt{\frac{1}{E-1}\sum_{e=1}^{E} |w_e - \bar{w}_e|^2}. \tag{19}$$

We normalize welfare so that the mean welfare per entrant of the No Policy scenario is equal to 1.

We conduct 50 rounds of the simulation, each with 13 years of draws d_{ikt}. We record the average of each of the previous statistics (E, E_t, W, $\bar{w}_e$, s_e) across the 50 rounds of simulation; this accounts for the randomness of the draws d_{ikt}.

We estimate the standard errors for the statistics E, E_t, W, $\bar{w}_e$, s_e using a nonparametric bootstrap. We randomly draw counties from the dataset with replacement to generate 250 independent panels of size equal to the actual sample size. These are the same datasets that we generated when bootstrapping the standard errors of the structural parameters. For each bootstrap sample, we run 50 simulations using the estimated parameters $\hat{\theta}$ and estimated probabilities of investment $\tilde{g}(\cdot)$ associated with the particular bootstrap draw, and then take the average of the statistics (E, E_t, W, $\bar{w}_e$, s_e) across the 50 rounds of simulation. The standard error is then formed by taking the standard deviation of the estimated statistics from each of the random samples.

The methodology for the counterfactual policy simulations is the same as that for the Base scenario replication, except we replace the indicators for the specified policy variables in G_{kt} with zero. For example, in the No RFS1 simulation, we set $RFS = 0$ for all observations.

For the No MTBE Ban scenario, we can only run the simulations for the pre-RFS period (1996-2004) because 2004 was the last year any state in our sample permitted the use of MTBE; we therefore never see cases in which there is no MTBE ban in combination with either RFS1 or RFS2, both of which were implemented after 2004. As a consequence, the effect of the MTBE ban is not identified after 2004.

One challenge in simulating alternate policy scenarios is that calculating the investment policy function $\tilde{g}(\cdot)$ relies on observing the corresponding state of the world in the data. Because entry in the simulation is random, we sometimes simulate states of the world that we do not observe in the data. We follow a rule to replace the missing value of $\tilde{g}(\cdot)$ for the simulated states of the world Ω^*_{kt} that we do not observe in the data. Many of the states of the world that are missing in the simulation occur because we draw an entrant in a county k that did not have an entrant in the data. Consequently, the first replacement rule is replace $\tilde{g}(N_{kt} = 1, G_{kt} = g, X_{kt} = x)$ with $\tilde{g}(N_{kt} = 0, G_{kt} = g, X_{kt} = x)$ when we do not observe $\Omega_{kt} = (N_{kt} = 1, G_{kt} = g, X_{kt} = x)$ in the data.

The second replacement rule to replace the missing value of $\tilde{g}(\cdot)$ for the simulated states of the world Ω^*_{kt} that we do not observe in the data is necessary because for some values of the state variables X_{kt}, we may not observe that value X_{kt} both when the policies G_{kt} are equal to zero and when the policies G_{kt} are not equal to zero. In the second rule, we find the state of the world $\Omega_{kt} = (N_{kt}, G_{kt}, X_{kt})$ where $\tilde{g}$ is defined, and where the variables in X_{kt} that have a statistically significant effect on the payoff from investing in building an ethanol plant and the policy variables G_{kt} match our simulated data.

The third and final replacement rule to replace the missing value of $\tilde{g}(\cdot)$ for the simulated states of the world Ω^*_{kt} that we do not observe in the data is to use the annual mean $\tilde{g}$ in place of the missing $\tilde{g}(\cdot)$.

Table 6 shows which replacement rule is used in each simulation. Each year there are a maximum of 2,610 potential entrants (3 in each of the 870 counties), for a total of 33,930 potential observations.[17] Virtually all replacements were made in Rule 1 or Rule 2. The No Policy simulation was the most challenging in this respect because there were relatively few years and counties among which to find replacements.

6.2 Goodness of Fit

To assess the goodness of fit of our model, we conduct a replication exercise in which we use our estimated model applied to the observed exogenous state variables to simulate (or predict) the data. We call the model predicted results our Base scenario.

[17] In practice, the total number of investor-year combinations is lower because once a potential investor invests, that investor exits the sample.

Table 7 compares the observed statistics E, E_t, W, $\bar{w}_e$, s_e in the data with their simulated values under the Base scenario. The Base scenario does a good job of replicating the observed number of entrants and their welfare: the simulated number of entrants in the Base scenario has a mean of 136, versus 132 in the data. The data and the Base scenario also have similar values for the mean welfare per entrant $\bar{w}_e$ and for total welfare E. Our model therefore does a fairly good job matching the statistics based on actual data.

6.3 Counterfactual Policy Scenarios

In addition to a simulated replication using the observed exogenous state variables (the Base scenario), we simulate investment under several counterfactual policy scenarios. These counterfactual scenarios include No RFS1, No RFS2, No Tax Credit, No MTBE Ban, and No Policy, and are summarized in Table 5.

For the No MTBE Ban scenario, we can only run the simulations for the pre-RFS period (1996-2004) because 2004 was the last year any state in our sample permitted the use of MTBE; we therefore never see cases in which there is no MTBE ban in combination with either RFS1 or RFS2, both of which were implemented after 2004. As a consequence, the effect of the MTBE ban is not identified after 2004.

In the first set of counterfactual policy scenarios, we remove each policy individually, and use two-sample t-tests to compare the results of these simulations to the those of the Base replication. As seen in the results in Table 8, removing the RFS2 significantly decreases the number of entrants compared to the Base, while removing RFS1 and the state tax credit have smaller but noticeable affects on the number of entrants as well. The removal of RFS2 also has the largest impact on the mean welfare per entrant $\bar{w}_e$ of the three policies, indicating RFS2 had the largest impact on entrant payoff of the four policies.

We also simulate the No Policy scenario, which removes all the policies, *MTBE ban*, *RFS 1*, *RFS 2*, and *Tax Credit*, that might promote investment in ethanol plants, and use two-sample t-tests to compare the results of each scenario to the those of the No Policy.

As seen in Table 8, there are two striking results that arise from comparing entrants and welfare in the Base and No Policy scenarios. First, the mean number of entrants with no policies is 37, versus 136 in the Base simulation. Together, the four policies led to most of the investment in plants over the 13 years of the simulation. The second important takeaway is that the mean welfare per entrant, $\bar{w}_e$, is significantly lower in the No Policy scenario than it is under the Base replication scenario. There is less entry because expected payoff from investment in an ethanol plant is much lower without the policies. The standard deviation of welfare per entrant is still large though; policy changes account for some, but not all, of the differences in profitability across space and time.

Table 9 shows the results of the No MTBE Ban scenario, as well as the Base, No Policy, and No Tax Credit scenarios for the pre-RFS period (1996-2004). We conducted these simulations through 2004 instead of through 2008 because it was not possible to identify states of the world with one of the RFS standards in place, but without the MTBE ban. In this period, there were 48 entrants in the Base replication (46 in the data: see Table 7), and 29 in the scenario with No MTBE Ban; this large difference is statistically significant. In this same time frame, there were 26 entrants in the No Policy scenario, and the difference between the No Policy scenario and the No MTBE Ban scenario is only marginally statistically significant.

In this same pre-RFS time period, the No Tax Credit scenario leads to fewer entrants that the Base replication, but this number is still more than the No MTBE Ban scenario. In aggregate, this results indicates that the MTBE Ban had a bigger effect on entry than the state tax credits in the pre-RFS era during which the effects of the two policies can be identified and compared.

We disaggregate these results by year in Table 10 to further explore the interactions among the policy effects. Viewing the simulated entrants by year is useful to begin to disentangle the effects of the MTBE Ban and the RFS. Figure 2 shows the cumulative number of entrants and the total cumulative welfare of entrants over time. Entry and total welfare of entrants increased faster in the later years of the analysis, particularly in the years during which the RFS2 was in effect (2007-2008).

As seen in Table 10, the No Tax Credit simulation yielded on average 9% fewer entrants per given year compared to the Base simulation. The impact was smaller in the earlier years of the simulation, when fewer states had policies in place. The No RFS1 simulation had a slightly larger impact on the number of entrants than the No Tax Credit simulation for the years when RFS1 was in effect (2005-2006), though the cumulative number of entrants was still greater under the No RFS1 scenario because it was in effect for fewer years.

The No RFS2 scenario led to a much more marked decrease in the number of entrants per year compared to the no RFS1 and No Tax Credit scenarios (Figure 2), though the number of entrants per year during the RFS2 period (2007-2008) was still greater than the beginning of our analysis period due to other favorable economic conditions (Table 10).

Though we can only identify the No MTBE Ban scenario in the pre-RFS era (before 2005), we find similar magnitude of impact on the number of entrants as the No RFS2 scenario, particularly as we get closer to 2005, when all the states in our analysis had banned MTBE.

In the No Policy scenario, entry was slow and relatively constant over time, ranging from 1.6 to 4.1 new plants each year. In the Base replication the number of entrants per year increased over time, with a maximum of 32.5 new plants in 2007 (the second to last year of the simulation).

Figure 3 shows how the mean welfare per entrant by year changed over time under each scenario. The lines for the No RFS1 and the No Tax Credit scenario closely track the Base replication, indicating that these policies had relatively small impacts on profitability for entrants. Nevertheless, both the No MTBE Ban and No RFS2 scenarios led to significantly lower welfare for entrants compared to the Base replication in respective the years when the MTBE ban and the RFS2 were in effect.

Welfare per entrant was lower in the pre-RFS era, which is why there were fewer entrants. The first states in our sample banned MTBE as early as 2000, when we see the welfare per entrant under the No MTBE Ban scenario drop significantly below that of the Base replication. During the period 2000-2004, which represents the period during which there were some MTBE bans but no RFS1 or RFS2, the MTBE ban accounted for 54% of the entrants in the period. Without the ban, there would have been 16 new plants instead of the 35 that entered in the Base scenario. The RFS2 had a larger impact in percentage and real terms. Nevertheless, the level of entry in the Base replication was higher in later years due to the combination of policy and market factors.

We disaggregate the results by each of the 10 Midwestern states in Table 11. Different states provide better environment for entering ethanol plants, as well as implementing the MTBE ban and offering tax credits at different times. Figure 4 shows how entry compares across states and policy scenarios. Each bar in the graph shows the number of entrants in the pre-RFS period (1996-2004) in black, and the number of entrants in the post-RFS period (2005-2008) in grey, for each state and each policy scenario. Figure 5 presents the mean welfare per entrant for each scenario by state, for the full period (left panel) and for the pre-RFS period (right panel).

There are noticeable differences across states in the total number of entrants, in the timing of the entrants, and in the relative impact of the different policy scenarios on entry. First, some states attract much more entry of ethanol plants than others under all scenarios. In particular, Iowa and Nebraska have the most entrants. The total number of entrants does not exactly correspond with the mean welfare per entrant, however (Figure 3). The mean welfare per entrant is high in these two states, but overall, entrants had higher welfare from entry in Indiana and South Dakota in the Base replication; South Dakota had fewer entrants because only part of the state is at all suitable for ethanol production.

The second important difference across states is that some states had relatively more entrants in the pre-RFS era than others. Nebraska, for example, had over half of its plants enter before 2005. Minnesota also experienced more entry in the pre-RFS era. Both these states implemented MTBE bans early (in 2000), and also had state tax credits for plants that gave them more favorable conditions for entrants.

Different policies had different impacts on different states. The number of pre-RFS entrants in the Base and No MTBE Ban scenarios is directly proportional to the number of years the MTBE ban was in effect in each state, indicating that this policy made a large contribution to industry growth in the region. Likewise, the No RFS2 scenario led to fewer entrants in all states, indicating that the RFS2 was a driver of industry growth in the last two years of our analysis.

The No Tax Credit scenario had more mixed results. All the states except Ohio, Iowa, and Illinois had tax credits available to entrants at some point during the analysis, though the year these policies were in effect varied across states.

7 Conclusions

In this paper, we develop and estimate a dynamic structural econometric model of the ethanol plant investment timing game to analyze the effects of government policy on the decisions of ethanol-producing firms to invest in building new ethanol plants in the Midwestern United States during the second US ethanol boom.

According to our results, government policies, particularly the ban on the use of the oxygenate MTBE as a gasoline additive, and the 2007 Renewable Fuel Standard (RFS2), have significant effects on ethanol investment payoffs and decisions. The intensity of corn production and private information shocks have significant effects on ethanol investment payoffs and decisions as well.

We use the estimated structural parameters to simulate counterfactual policy scenarios to disentangle the impacts of state and national policies on the timing and location of investment in the industry. We find that, of the policies analyzed, the MTBE ban and the RFS2 led to most of the ethanol plant investment during this time period. There are noticeable differences across states in the total number of entrants, in the timing of the entrants, and in the relative impact of the different policy scenarios on entry.

One possible reason the MTBE ban was effective in inducing investment in building ethanol plants is that it increased the demand for ethanol as an oxygenate in place of MTBE. Similarly, one possible reason the RFS2 was effective in inducing investment in building ethanol plants is that it increased demand for ethanol by mandating an expansion in ethanol consumption. Previous studies have shown that the demand for ethanol is highly elastic

(Irwin and Good, 2017; Yi, Lin Lawell and Thome, 2018). Our results suggest that policies that increase the demand for ethanol have the potential for inducing investment in building ethanol plants.

Both the MTBE ban and the Renewable Fuel Standard can function as implicit blending mandates (de Gorter and Just, 2010; Anderson and Elzinga, 2014). Whenever unpriced emissions are the sole market failure, however, a carbon tax or cap and trade program is more likely to achieve the first-best (Pigou, 1920; Coase, 1960), while fuel mandates are unable to replicate the first-best solution (Helfand, 1992; Holland, Knittel and Hughes, 2009; Lapan and Moschini, 2012). Lade and Lin Lawell (2018) show that when renewable fuel mandates are combined with a cost containment mechanism such as a credit window price, the efficiency of the mandate can increase substantially. Thus, while the MTBE ban and the Renewable Fuel Standard were effective in inducing investment in building ethanol plants, it is possible to increase their efficiency by combining them with cost containment mechanisms or by using a market-based instrument instead. We hope to explore these possibilities in future work.

References

Agarwal, N., I. Ashlagi, M. Rees, P. Somaini, and D. Waldinger. (2018). An empirical framework for sequential assignment: The allocation of deceased donor kidneys. Working paper, MIT, Stanford University, and University of Toledo.

Aguirregabiria, V., and A. Luengo. (2016). A microeconometric dynamic structural model of copper mining decisions. Working paper.

Aguirregabiria, V., and P. Mira. (2007). Sequential estimation of dynamic discrete games. *Econometrica*, 75 (1), 1-53.

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.

Bajari, P., C.L. Benkard, and J. Levin. (2007). Estimating dynamic models of imperfect competition. *Econometrica*, 75 (5), 1331-1370.

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.

Barwick, P. J., M. Kalouptsidi, and N. Zahur. (2018). China's industrial policy, excess capacity, and fragmentation. Working paper, Cornell University and Harvard University.

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

Carroll, C.L., C.A. Carter, R.E. Goodhue, and C.-Y.C. Lin Lawell. (2018a). Supply chain externalities and agricultural disease. Working paper, Cornell University.

Carroll, C.L., C.A. Carter, R.E. Goodhue, and C.-Y.C. Lin Lawell. (2018b). The economics of decision-making for crop disease control. Working paper, Cornell University.

Carroll, C.L., C.A. Carter, R.E. Goodhue, and C.-Y.C. Lin Lawell. (forthcoming). Crop disease and agricultural productivity: Evidence from a dynamic structural model of Verticillium wilt management. In W. Schlenker (Ed.), *Understanding Productivity Growth in Agriculture*. Chicago: University of Chicago Press.

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.

Cook, J.A., and C.-Y.C. Lin Lawell. (2018). Wind turbine shutdowns and upgrades in Denmark: Timing decisions and the impact of government policy. Working paper, Cornell University.

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.

Dixit, A.K., and R.S. Pindyck. (1994). *Investment Under Uncertainty*. Princeton, NJ: Princeton University Press.

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

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.

Feger, F., N. Pavanini, N., and D. Radulescu. (2017). Welfare and redistribution in residential electricity markets with solar power. Working paper.

Fowlie, M., Reguant, M., and Ryan, S.P. (2016). Market-based emissions regulation and industry dynamics. *Journal of Political Economy*, 124 (1), 249-302.

Fudenberg, D., and J. Tirole. (1998). *Game Theory*. Cambridge: MIT Press.

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.

Gerarden, T. (2018). Demanding innovation: The impact of consumer subsidies on solar panel production costs. Working paper, Cornell University.

Gillingham, K., F. Iskhakov, A. Munk-Nielsen, J. Rust, and B. Schjerning. (2016). A dynamic model of vehicle ownership, type choice, and usage. Working paper.

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.

Harsanyi, J. (1973). Games with randomly disturbed payoffs. *International Journal of Game Theory*, 2, 1-23.

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.

Hotz, V.J., R.A. Miller, S. Sanders, and J. Smith. (1994). A simulation estimator for dynamic models of discrete choice. *Review of Economic Studies*, 61, 265-289.

Huang, L., and M.D. Smith. (2014). The dynamic efficiency costs of common-pool resource exploitation. *American Economic Review*, 104 (12), 4071-4103.

Irwin, S., and D. Good. (2017). On the value of ethanol in the gasoline blend. *farmdoc daily*, 7 (48). Department of Agricultural and Consumer Economics, University of Illinois at Urbana-Champaign, 15 March 2017. URL: `https://farmdocdaily.illinois.edu/2017/03/on-the-value-of-ethanol-in-the-gasoline-blend.html`

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

Jha, A. (2018). Dynamic regulatory distortions: Coal procurement at U.S. power plants. Working paper, Carnegie Mellon University.

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.

Kalouptsidi, M. (2018). Detection and impact of industrial subsidies: The case of Chinese shipbuilding. *Review of Economic Studies*, 85 (2), 1111-1158.

Kheiravar, K.H., C.-Y.C. Lin Lawell, and A.M. Jaffe. (2018). A structural econometric model of the dynamic game between petroleum producers in the world petroleum market. Working paper, Cornell University.

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.

Langer, A. and D. Lemoine. (2018). Designing dynamic subsidies to spur adoption of new technologies. NBER Working Paper No. 24310.

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.

Leyden, B.T. (2018). There's an app (update) for that: Understanding product updating under digitization. Working paper, Cornell University.

Li, S., and C.D. Wei. (2013). Green stimulus: A dynamic discrete analysis of vehicle scrappage programs. Working paper, Cornell University.

Lim, C.S.H., and A. Yurukoglu. (2018). Dynamic natural monopoly regulation: Time inconsistency, moral hazard, and political environments. *Journal of Political Economy*, 126 (1), 263-312. Lin, C.-Y.C. (2013). Strategic decision-making with information and extraction externalities: A structural model of the multi-stage investment timing game in offshore petroleum production. *Review of Economics and Statistics*, 95 (5), 1601-1621.

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.

Ma, X., C.-Y.C. Lin Lawell, and S. Rozelle. (2018). Estimating peer effects: A structural econometric model using a field experiment of a health promotion program in rural China. Working paper, Cornell University.

Mahajan, A., and A. Tarozzi. (2011). Time inconsistency, expectations and technology adoption: The case of insecticide treated nets. Working paper, University of California at Berkeley.

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`.

Pakes, A., M. Ostrovsky, and S. Berry. (2007). Simple estimators for the parameters of discrete dynamic games (with entry and exit examples). *RAND Journal of Economics*, 38 (2), 373-399.

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.

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

Rapson, D. (2014). Durable goods and long-run electricity demand: Evidence from air conditioner purchase behavior. of Environmental Economics and Management, 68 (1), 141160.

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.

Rojas Valdés, R.I., C.-Y.C. Lin Lawell, and J.E. Taylor. (2018a). Neighborhood effects in migration decisions in rural Mexico. Working paper, Cornell University.

Rojas Valdés, R.I., C.-Y.C. Lin Lawell, and J.E. Taylor. (2018b). The dynamic migration game: A structural econometric model and application to rural Mexico. Working paper, Cornell University.

Rothwell, G., and J. Rust. (1997). On the optimal lifetime of nuclear power plants. *Journal of Business and Economic Statistics*, 15 (2), Structural Estimation in Applied Microeconomics, 195-208.

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

Rust, J. (1988). Maximum likelihood estimation of discrete control processes. *SIAM Journal on Control and Optimization*, 26 (5), 10061024.

Ryan, S.P. (2012). The costs of environmental regulation in a concentrated industry. *Econometrica*, 80 (3), 1019-1061.

Sambucci, O., C.-Y.C. Lin Lawell, and T.J. Lybbert. (2018). Pesticide spraying and disease forecasts: A dynamic structural econometric model of grape growers in California. Working paper, Cornell University.

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.

Scott, P.T. (2013). Dynamic discrete choice estimation of agricultural land use. Working paper, New York University.

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.

Stokey, N., R. Lucas, and E. Prescott. (1989). *Recursive Methods in Economic Dynamics*. Cambridge, MA: Harvard University Press.

Thome, K.E., and C.-Y.C. Lin Lawell. (2018). Corn ethanol in the Midwestern United States: Local competition, entry, and agglomeration. 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.

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.

Zakerinia, S., and C.-Y.C. Lin Lawell. (2018). Climate change policy: Dynamics, strategy, and the Kyoto Protocol. Working paper, Cornell University.

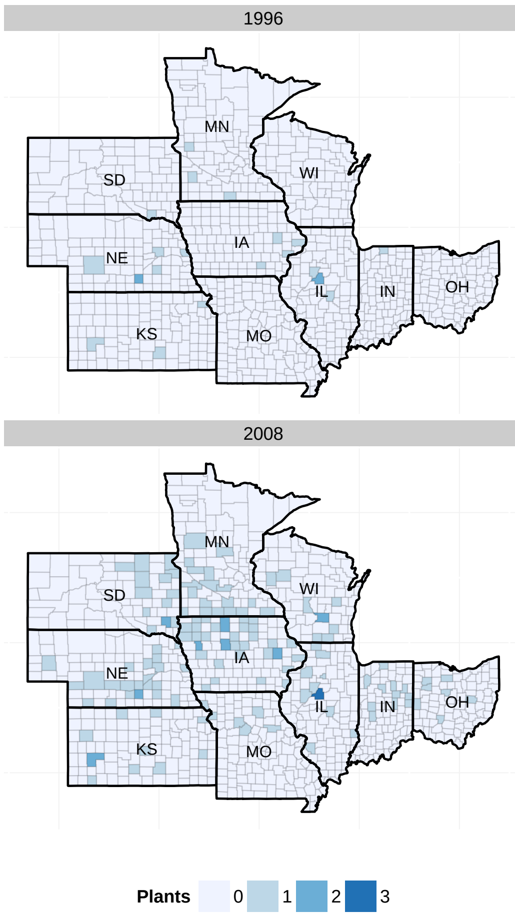

Figure 1: Number of operational ethanol plants by county in the Midwestern United States

Table 1: Bin design of variables for structural estimation

	Base Bins		Alternate Bins		
Variable	Bin Design	Break	Bin Design	Break 1	Break 2
Cow Density (head/acre)	Bottom two thirds and top third	0.103	Middle Bin is 1.5 Std. Dev. around Mean	0.048	0.124
Corn Intensity	Equal sizes	0.175	Middle Bin is 1.5 Std. Dev. around Mean	0.078	0.191
Ethanol Price ($/gal)	Equal sizes	1.630	Middle Bin is middle 5 years	1.51	1.91
Gasoline Price ($/gal)	Equal sizes	1.110	Bottom third and top two thirds	1	
Output Price Indicator	High if both ethanol and gasoline prices are high				
Alternate Corn Price ($/bushel)	Equal sizes	3.010	Middle Bin is 1.5 Std. Dev. around Mean	2.317	3.32
Corn Price ($/bushel)	Bottom third and top two thirds	2.340	Middle Bin is 2 Std. Dev. around Mean	5.48	8.88
Natural Gas Price ($/1000ft3)	Equal sizes	6.810	Middle Bin is 2 Std. Dev. around Mean	4.519	8.349
Electricity Price (cents/KwH)	Equal sizes	5.130	Middle Bin is 2 Std. Dev. around Mean	4.702	5.741
Energy Input Price Indicator	High if both electricity and natural gas prices are high				

Note: Corn intensity is defined as the corn acreage divided by the total area of the county.

Table 2: Summary statistics for discretized variables used in structural estimation

Variable	Base Bins		Alternate Bins		Spatial Resolution
	Mean	Std. Dev.	Mean	Std. Dev.	
New Plant	0.004	0.063	0.004	0.063	county
Tax Credit	0.341	0.474	0.341	0.474	state policy
MTBE Ban	0.476	0.499	0.475	0.499	state policy
RFS I	0.153	0.360	0.153	0.360	national policy
RFS II	0.151	0.358	0.151	0.358	national policy
Cow Density	0.330	0.470	0.943	0.760	district (USDA definition)
Corn Intensity	0.494	0.500	0.917	0.669	county
Corn Price	0.677	0.468	0.918	0.712	state
Alternate Corn Price	0.513	0.500			state
Soy Price			1.093	0.596	state
Output Price Indicator	0.648	0.478			state
Ethanol Price	0.535	0.499	0.917	0.728	national
Gasoline Price	0.493	0.500	0.380	0.485	state
Energy Input Price Indicator	0.797	0.402			state
Natural Gas Price	0.492	0.500	0.945	0.649	state
Electricity Price	0.499	0.500	0.971	0.545	state
Metro Area	0.283	0.450	0.010	0.099	county
Existing Plant	0.036	0.185	0.037	0.188	county
Biodiesel Plant	0.010	0.100	0.285	0.452	county
Number of Observations	33,307		33,307		
Number of Counties	870		870		

Table 3: Results of structural model

	Base Model	Alternate price specifications				
	(i)	(ii)	(iii)	(iv)	(v)	(vi)
Coefficients in the investment payoff on:						
Tax Credit	0.209	0.206	0.179	0.16	0.216	0.26
	(0.147)	(0.147)	(0.157)	(0.154)	(0.154)	(0.178)
MTBE Ban	0.814**	1.022***	0.837**	0.936**	0.907*	0.956*
	(0.293)	(0.303)	(0.305)	(0.299)	(0.372)	(0.323)
RFS 1	0.085	0.05	0.168	0.181	0.166	0.142
	(0.242)	(0.214)	(0.283)	(0.313)	(0.26)	(0.279)
RFS 2	0.727**	0.658**	0.786*	0.946**	0.816**	0.965***
	(0.256)	(0.231)	(0.32)	(0.338)	(0.309)	(0.27)
Cow Density	0.189	0.184	0.206	0.28‡	0.22	0.229
	(0.149)	(0.136)	(0.155)	(0.16)	(0.129)	(0.162)
Corn Intensity	1.012***	0.976***	0.962***	1.193***	0.986***	1.217***
	(0.181)	(0.163)	(0.201)	(0.198)	(0.213)	(0.22)
Energy Output Indicator	−0.423‡	−0.573‡			-0.542	-0.429
	(0.246)	(0.307)			(0.348)	(0.334)
Ethanol Price				-0.376		
				(0.364)		
Gasoline price			-0.289	-0.096		
			(0.286)	(0.245)		
Corn Price	-0.074	-0.071	-0.08	-0.085	-0.167	-0.183
	(0.265)	(0.197)	(0.239)	(0.205)	(0.259)	(0.231)
Energy Input Indicator		0.753*	0.517		0.67	
		(0.354)	(0.41)		(0.444)	
Natural Gas Price	0.374			0.436		0.383
	(0.275)			(0.404)		(0.349)
Electricity Price				0.036		
				(0.179)		
Existing Plant	0.034	0.021	-0.237	0.042	-0.129	0.039
	(0.279)	(0.286)	(0.29)	(0.311)	(0.268)	(0.307)
Constant	-4.97***	-5.164***	-5.087***	-6.108***	-5.042***	-5.962***
	(0.411)	(0.372)	(0.512)	(0.403)	(0.413)	(0.506)
σ	0.648***	0.612***	0.61***	0.786***	0.606***	0.776***
	(0.042)	(0.039)	(0.048)	(0.043)	(0.073)	(0.051)
Number of Observations	33,307	33,307	33,307	33,307	33,307	33,307
Number of Counties	870	870	870	870	870	870

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

Table 4: Results of structural model with alternate variable and bin specifications

	Base Model	Additional Covariates	Alternate (More) Bins	
	(i)	(vii)	(viii)	(ix)
Coefficients in the investment payoff on:				
Tax Credit	0.209	0.123	0.109	0.394
	(0.147)	(0.135)	(0.247)	(0.398)
MTBE Ban	0.814**	1.044***	0.502‡	1.014***
	(0.293)	(0.296)	(0.268)	(0.284)
RFS 1	0.085	0.044	1.287***	1.674***
	(0.242)	(0.209)	(0.295)	(0.403)
RFS 2	0.727**	0.651*	2.343***	1.869***
	(0.256)	(0.268)	(0.322)	(0.266)
Cow Density	0.189	0.225‡	0.708***	0.812***
	(0.149)	(0.131)	(0.159)	(0.13)
Corn Intensity	1.012***	0.965***	0.209	0.315*
	(0.181)	(0.168)	(0.173)	(0.131)
Energy Output Price Indicator	-0.423	-0.586*		
	(0.246)	(0.281)		
Ethanol Price			-0.518	-1.916**
			(0.646)	(0.636)
Gasoline Price			2.168***	2.546***
			(0.551)	(0.613)
Corn Price	-0.074	-0.071	-0.439	0.089
	(0.265)	(0.216)	(0.34)	(0.266)
Soy Price			-0.493	0.67
			(0.59)	(0.758)
Energy Input Price Indicator		0.792*		
		(0.382)		
Natural Gas Price	0.374		-1.549*	-1.104*
	(0.275)		(0.69)	(0.474)
Electricity Price			-0.179	
			(0.253)	
Metro Area		-0.244	-0.564	-0.369
		(0.2)	(0.569)	(0.589)
Existing Plant	0.034	-0.123	0.135	-0.017
	(0.279)	(0.26)	(0.363)	(0.347)
Biodiesel Plant		-0.06	0.033	0.023
		(0.48)	(0.084)	(0.074)
Constant	-4.97***	-5.08***	-5.591***	-6.583***
	(0.411)	(0.287)	(0.607)	(0.587)
σ	0.648***	0.609***	0.997***	0.77***
	(0.042)	(0.046)	(0.083)	(0.092)
Number of Observations	33,307	33,307	33,307	33,307
Number of Counties	870	870	870	870

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

Table 5: Counterfactual Scenarios

Counterfactual Scenario	Description
Base Scenario	Replication with observed data
No RFS1	Remove RFS1 (set *RFS1*=0)
No RFS2	Remove RFS1 (set *RFS2*=0)
No Tax Credit	Remove state tax credit (set *Tax Credit*=0)
No MTBE Ban	Remove MTBE ban (set *MTBE ban*=0) [Pre-RFS (1996-2004) only]
No Policy	Remove all policies (set all G_{kt} variables=0)

Table 6: Replacement rules followed in counterfactual simulations for missing entry probabilities $\tilde{g}(\cdot)$

		Replacement Rule Followed:		
Counterfactural Scenario	Number Missing	Set *Existing plant*=0	Match policy and significant state variables	Use annual mean $\tilde{g}$
Base Scenario	48.5	48.4	0.1	0.0
No RFS1	66.3	65.5	0.8	0.0
No RFS2	101.2	99.5	1.7	0.0
No Tax Credit	827.3	88.7	738.5	0.0
No MTBE Ban (1996-2004)	380.9	168.1	212.8	0.0
No Policy	4209.1	427.7	3781.2	0.2

Notes: The replacement rules are used to replace the missing value of $\tilde{g}(\cdot)$ for the simulated states of the world Ω^*_{kt} that we do not observe in the data.

Table 7: Number of entrants and welfare in data and Base scenario

Full Period			Welfare per Entrant	
	Number of Entrants	Total Welfare of Entrants	Mean	Std. Dev.
Data	132	273.28	2.07	0.64
Base Scenario	135.92	278.21	2.05	0.704
	(14.97)	(31.62)	(0.15)	(0.04)
1996-2004			Welfare per Entrant	
	Number of Entrants	Total Welfare of Entrants	Mean	Std. Dev.
Data	46	65.68	1.43	0.596
Base Scenario	47.60	64.01	1.3449	0.6958
	(14.38)	(32.41)	(0.154)	(0.0649)

Notes: For the Base scenario, the reported statistics are averages over 50 simulations. Standard errors are in parentheses, and are calculated from using the parameter estimates from each of the 250 bootstrap samples. For each of the 250 bootstrap samples, 50 simulations are run using the parameter estimates from that bootstrap sample. Standard errors for a statistic is the standard deviation of the respective statistics over all 250 bootstrap samples.

Table 8: Number of entrants and welfare under counterfactual policy scenarios: Full Period

	Base Scenario	No RFS1	No RFS2	No Tax Credit	No Policy
Number of Entrants	135.9	131.6	91.8	123.2	36.6
	(15.0)	(17.1)	(17.8)	(15.7)	(17.8)
Total Welfare of All Entrants	278.2	267.8	157.0	246.6	36.6
	(31.6)	(35.4)	(36.7)	(34.0)	(42.6)
Mean of Welfare per Entrant	2.05	2.03	1.71	2.00	1.00
	(0.15)	(0.14)	(0.14)	(0.15)	(0.18)
Std. Dev. of Welfare per Entrant	0.70	0.72	0.65	0.72	0.60
	(0.04)	(0.04)	(0.04)	(0.04)	(0.07)
Difference between this scenario and Base scenario					
Number of Entrants		-4.3**	-44.1***	-12.7***	-99.3***
Total Welfare of All Entrants		-10.4	-121.2***	-31.6***	-241.6***
Mean of Welfare per Entrant		-0.02	-0.34***	-0.05***	-1.05***
Std. Dev. of Welfare per Entrant		0.02**	-0.05***	0.02**	-0.10***
Difference between this scenario and No Policy scenario					
Number of Entrants	99.3***	95.0***	55.2***	86.6***	
Total Welfare of All Entrants	241.6***	231.2***	120.4***	210.0***	
Mean of Welfare per Entrant	1.05***	1.03***	0.71***	1.00***	
Std. Dev. of Welfare per Entrant	0.10***	0.12***	0.05***	0.12***	

Notes: For each scenario, the reported statistics are averages over 50 simulations. Standard errors are in parentheses, and are calculated from using the parameter estimates from each of the 250 bootstrap samples. For each of the 250 bootstrap samples, 50 simulations are run using the parameter estimates from that bootstrap sample. Standard errors for a statistic is the standard deviation of the respective statistics over all 250 bootstrap samples. Significance codes for two-sample t-tests of difference between scenarios: *** p<0.001, ** p<0.01, * p<0.05

Table 9: Number of entrants and welfare under counterfactual policy scenarios: Pre-RFS period (1996-2004)

	Base Scenario (to 2005)	No Tax Credit	No MTBE Ban	No Policy
Number of Entrants	47.6	43.0	28.9	26.1
	(14.4)	(14.4)	(16.0)	(15.6)
Total Welfare of All Entrants	64.0	54.6	24.6	20.3
	(32.4)	(33.5)	(35.7)	(36.2)
Mean of Welfare per Entrant	1.34	1.27	0.85	0.78
	(0.15)	(0.15)	(0.21)	(0.19)
Std. Dev. of Welfare per Entrant	0.70	0.69	0.56	0.55
	(0.06)	(0.07)	(0.06)	(0.06)
	Difference between this scenario and Base scenario			
Number of Entrants		-4.6***	-18.7***	-21.5***
Total Welfare of All Entrants		-9.4*	-39.4***	-43.7***
Mean of Welfare per Entrant		-0.07***	-0.49***	-0.56***
Std. Dev. of Welfare per Entrant		-0.01	-0.14***	-0.15***
	Difference between this scenario and No Policy scenario			
Entrants	21.5***	16.9***	2.8*	
Total Welfare of All Entrants	43.7***	34.3***	4.3	
Mean of Welfare per Entrant	0.56***	0.49***	0.07***	
Std. Dev. of Welfare per Entrant	0.15***	0.14***	0.01	

Notes: For each scenario, the reported statistics are averages over 50 simulations. Standard errors are in parentheses, and are calculated from using the parameter estimates from each of the 250 bootstrap samples. For each of the 250 bootstrap samples, 50 simulations are run using the parameter estimates from that bootstrap sample. Standard errors for a statistic is the standard deviation of the respective statistics over all 250 bootstrap samples. Significance codes for two-sample t-tests of difference between scenarios: ***p<0.001, **p<0.01, *p<0.05

Table 10: Number of entrants and mean welfare per entrant by year

Number of Entrants	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008
Base	1.7	3.4	3.6	4.4	4.2	6.8	7.2	6.7	9.8	13.7	11.2	32.5	30.9
	(0.7)	(1.4)	(3.0)	(6.8)	(1.8)	(1.8)	(2.2)	(1.8)	(2.8)	(3.7)	(3.2)	(5.4)	(4.9)
No RFS1	1.7	3.4	3.6	4.4	4.2	6.8	7.3	6.7	9.8	10.9	9.4	32.6	31.0
	(0.7)	(1.4)	(3.1)	(6.8)	(1.8)	(1.8)	(2.1)	(1.7)	(2.8)	(3.1)	(3.9)	(5.4)	(5.0)
No RFS2	1.7	3.4	3.6	4.4	4.2	6.8	7.3	6.7	9.8	13.8	11.2	9.7	9.4
	(0.7)	(1.4)	(3.1)	(6.8)	(1.8)	(1.8)	(2.1)	(1.7)	(2.8)	(3.7)	(3.2)	(4.2)	(4.0)
No Tax Credit	1.6	3.1	3.4	4.1	3.8	5.8	6.7	5.8	8.7	12.6	10.0	29.5	28.1
	(0.7)	(1.4)	(2.9)	(6.6)	(1.6)	(1.5)	(2.2)	(1.6)	(2.6)	(3.4)	(2.9)	(5.2)	(4.9)
No MTBE	1.7	3.3	3.6	4.4	2.3	3.1	3.5	3.5	3.5				
	(0.7)	(1.4)	(3.1)	(6.8)	(1.9)	(1.3)	(1.3)	(1.1)	(1.5)				
No Policy	1.6	3.1	3.4	4.1	2.1	2.7	3.2	3.0	3.0	3.2	2.4	2.5	2.4
	(0.7)	(1.4)	(2.9)	(6.6)	(1.8)	(1.1)	(1.2)	(0.9)	(1.3)	(1.4)	(0.9)	(1.0)	(1.0)
Mean Welfare per Entrant	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008
Base	-0.29	0.37	0.57	0.77	1.10	1.52	1.59	1.66	1.86	2.04	2.04	2.51	2.56
	(0.33)	(0.24)	(0.26)	(0.30)	(0.22)	(0.19)	(0.22)	(0.17)	(0.16)	(0.19)	(0.23)	(0.21)	(0.20)
No RFS1	-0.28	0.36	0.56	0.78	1.11	1.52	1.60	1.66	1.86	1.95	1.96	2.51	2.56
	(0.33)	(0.24)	(0.27)	(0.30)	(0.22)	(0.19)	(0.21)	(0.17)	(0.16)	(0.16)	(0.24)	(0.21)	(0.21)
No RFS2	-0.28	0.36	0.56	0.78	1.11	1.52	1.60	1.66	1.86	2.04	2.04	2.05	2.12
	(0.33)	(0.24)	(0.27)	(0.30)	(0.22)	(0.19)	(0.21)	(0.17)	(0.16)	(0.19)	(0.23)	(0.23)	(0.22)
No Tax Credit	-0.35	0.30	0.50	0.72	1.02	1.44	1.55	1.60	1.81	2.01	2.00	2.48	2.53
	(0.31)	(0.23)	(0.25)	(0.29)	(0.19)	(0.19)	(0.22)	(0.18)	(0.16)	(0.19)	(0.22)	(0.21)	(0.21)
No MTBE Ban	-0.28	0.36	0.57	0.78	0.60	0.94	1.12	1.27	1.35				
	(0.32)	(0.25)	(0.26)	(0.30)	(0.26)	(0.25)	(0.19)	(0.18)	(0.21)				
No Policy	-0.35	0.30	0.50	0.72	0.55	0.87	1.08	1.22	1.29	1.43	1.44	1.56	1.66
	(0.31)	(0.23)	(0.25)	(0.29)	(0.24)	(0.23)	(0.18)	(0.18)	(0.19)	(0.19)	(0.23)	(0.22)	(0.21)

Notes: For each scenario, the reported statistics are averages over 50 simulations. Standard errors are in parentheses, and are calculated from using the parameter estimates from each of the 250 bootstrap samples. For each of the 250 bootstrap samples, 50 simulations are run using the parameter estimates from that bootstrap sample. Standard errors for a statistic is the standard deviation of the respective statistics over all 250 bootstrap samples.

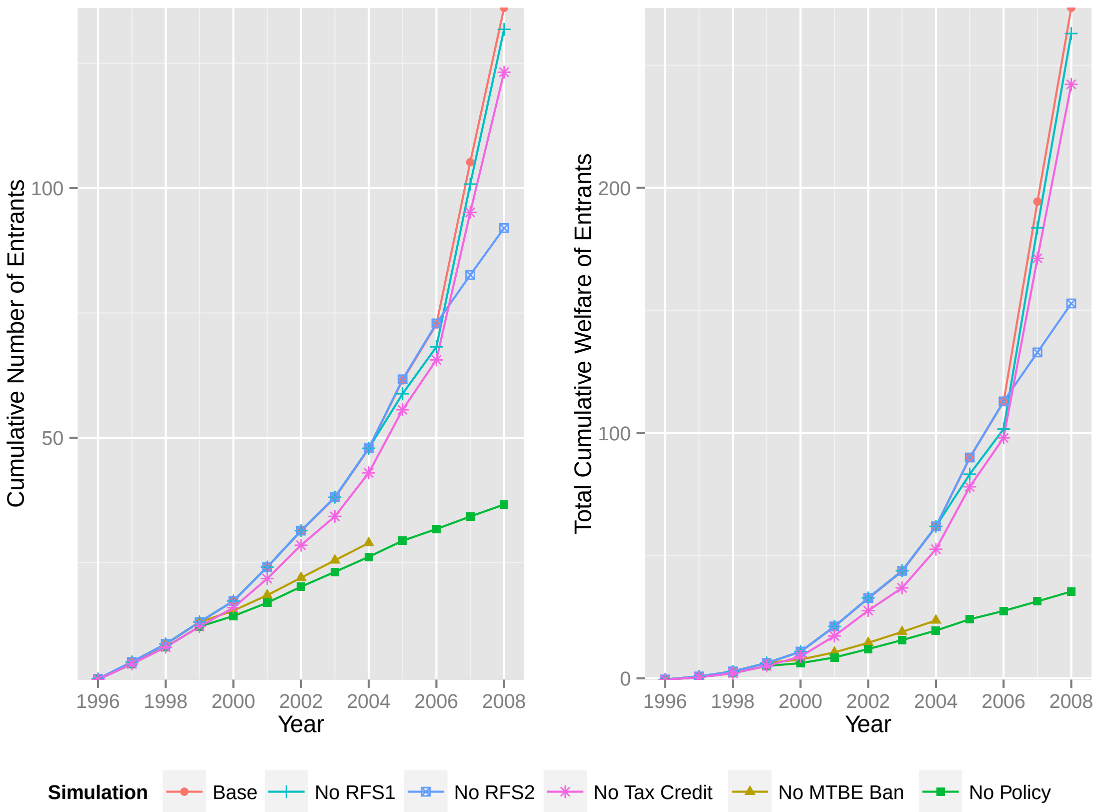

Figure 2: Cumulative number of entrants and total cumulative welfare of entrants under different policy scenarios over time

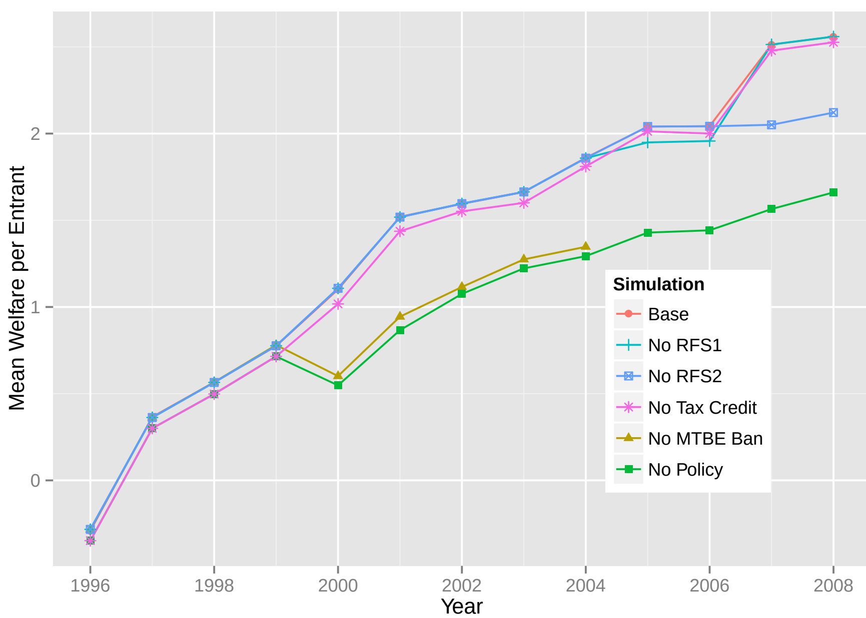

Figure 3: Mean welfare per entrant by year under different policy scenarios

Table 11: Number of entrants and mean welfare per entrant by state in full and pre-RFS periods

Number of Entrants		IL	IN	IA	KS	MN	MO	NE	OH	SD	WI
Full Period	Base	16.2	16.8	23.6	8.7	14.3	8.1	19.9	8.4	12.2	7.6
		(2.7)	(3.3)	(4.2)	(1.8)	(2.8)	(1.6)	(3.4)	(1.6)	(2.9)	(1.7)
	No RFS1	15.6	16.2	22.9	8.4	14.0	7.8	19.5	8.1	11.9	7.2
		(2.9)	(3.2)	(4.5)	(1.9)	(2.9)	(1.5)	(3.7)	(1.6)	(3.0)	(1.7)
	No RFS2	10.2	10.3	16.7	5.6	10.5	5.0	15.0	5.3	8.0	5.2
		(2.4)	(2.5)	(3.9)	(1.7)	(2.8)	(1.4)	(3.2)	(1.3)	(2.9)	(1.5)
	No Tax Credit	16.2	14.3	23.6	6.9	13.5	6.4	17.7	8.4	9.5	6.7
		(2.7)	(2.6)	(4.2)	(1.8)	(2.6)	(1.5)	(3.4)	(1.6)	(2.7)	(1.6)
	No Policy	5.1	4.9	6.1	2.2	3.8	2.2	4.7	3.1	2.5	2.1
		(1.8)	(2.3)	(2.6)	(1.7)	(2.6)	(1.4)	(2.0)	(1.1)	(2.7)	(1.2)
Pre-RFS (1996-2004)	Base	4.2	4.3	9.7	2.4	6.6	1.8	9.9	2.2	4.1	2.3
		(1.3)	(2.1)	(2.6)	(1.3)	(2.5)	(1.0)	(2.5)	(0.8)	(2.4)	(1.1)
	No Tax Credit	4.2	4.3	9.7	2.0	5.7	1.6	7.7	2.2	3.7	1.9
		(1.4)	(2.0)	(2.6)	(1.4)	(2.3)	(1.2)	(2.0)	(0.8)	(2.5)	(1.0)
	No MTBE Ban	3.3	3.5	4.3	1.9	3.4	1.8	4.7	2.2	2.0	1.9
		(1.3)	(2.1)	(2.2)	(1.4)	(2.7)	(1.0)	(2.1)	(0.9)	(2.6)	(1.1)
	No Policy	3.4	3.5	4.3	1.6	2.8	1.6	3.5	2.2	1.8	1.5
		(1.3)	(2.1)	(2.1)	(1.4)	(2.4)	(1.2)	(1.6)	(0.8)	(2.6)	(1.1)
Mean Welfare per Entrant		IL	IN	IA	KS	MN	MO	NE	OH	SD	WI
Full Period	Base	2.05	2.15	2.11	1.80	1.99	1.89	2.05	1.89	2.19	2.02
		(0.19)	(0.18)	(0.14)	(0.17)	(0.13)	(0.21)	(0.15)	(0.21)	(0.17)	(0.17)
	No RFS1	2.03	2.14	2.10	1.79	1.97	1.88	2.05	1.87	2.18	1.99
		(0.19)	(0.17)	(0.14)	(0.16)	(0.13)	(0.21)	(0.14)	(0.20)	(0.17)	(0.16)
	No RFS2	1.64	1.74	1.81	1.41	1.70	1.45	1.81	1.46	1.84	1.69
		(0.18)	(0.16)	(0.13)	(0.17)	(0.13)	(0.20)	(0.15)	(0.18)	(0.17)	(0.19)
	No Tax Credit	2.05	2.03	2.11	1.67	2.00	1.75	2.04	1.89	2.03	1.99
		(0.20)	(0.15)	(0.14)	(0.17)	(0.13)	(0.20)	(0.16)	(0.21)	(0.15)	(0.18)
	No Policy	1.01	1.09	1.12	0.60	1.04	0.70	1.08	0.94	0.98	0.93
		(0.18)	(0.17)	(0.18)	(0.17)	(0.20)	(0.19)	(0.20)	(0.19)	(0.23)	(0.21)
Pre-RFS (1996-2004)	Base	1.02	1.14	1.53	0.81	1.46	0.62	1.64	0.74	1.50	1.12
		(0.15)	(0.16)	(0.15)	(0.19)	(0.16)	(0.24)	(0.18)	(0.20)	(0.19)	(0.21)
	No Tax Credit	1.01	1.14	1.53	0.66	1.40	0.46	1.49	0.74	1.40	0.97
		(0.15)	(0.16)	(0.15)	(0.18)	(0.15)	(0.19)	(0.18)	(0.19)	(0.18)	(0.19)
	No MTBE Ban	0.72	0.88	0.91	0.48	0.92	0.62	1.04	0.74	0.84	0.85
		(0.17)	(0.20)	(0.20)	(0.22)	(0.26)	(0.24)	(0.26)	(0.21)	(0.28)	(0.26)
	No Policy	0.71	0.88	0.90	0.35	0.82	0.46	0.89	0.74	0.73	0.70
		(0.17)	(0.20)	(0.20)	(0.20)	(0.23)	(0.19)	(0.22)	(0.19)	(0.25)	(0.22)

Notes: For each scenario, the reported statistics are averages over 50 simulations. Standard errors are in parentheses, and are calculated from using the parameter estimates from each of the 250 bootstrap samples. For each of the 250 bootstrap samples, 50 simulations are run using the parameter estimates from that bootstrap sample. Standard errors for a statistic is the standard deviation of the respective statistics over all 250 bootstrap samples.

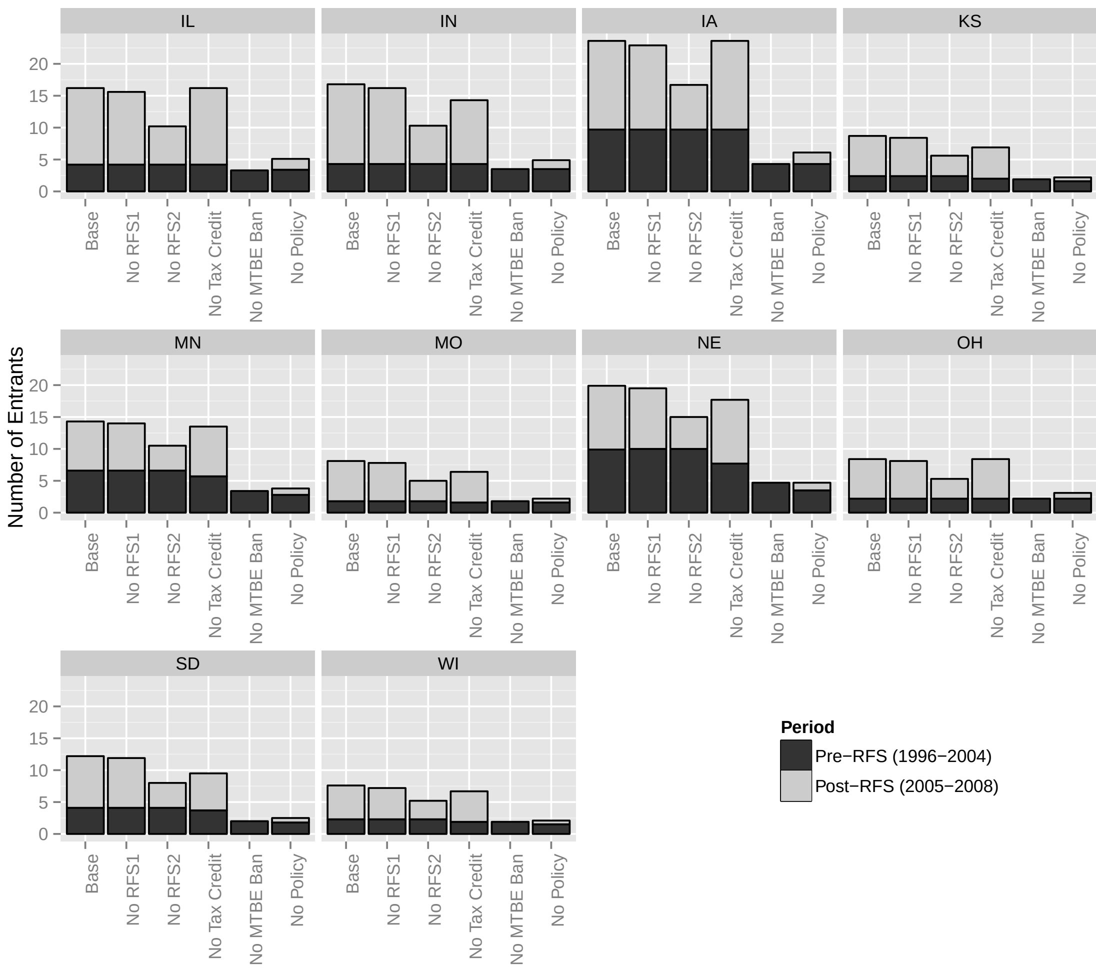

Figure 4: Number of entrants by state under different policy scenarios

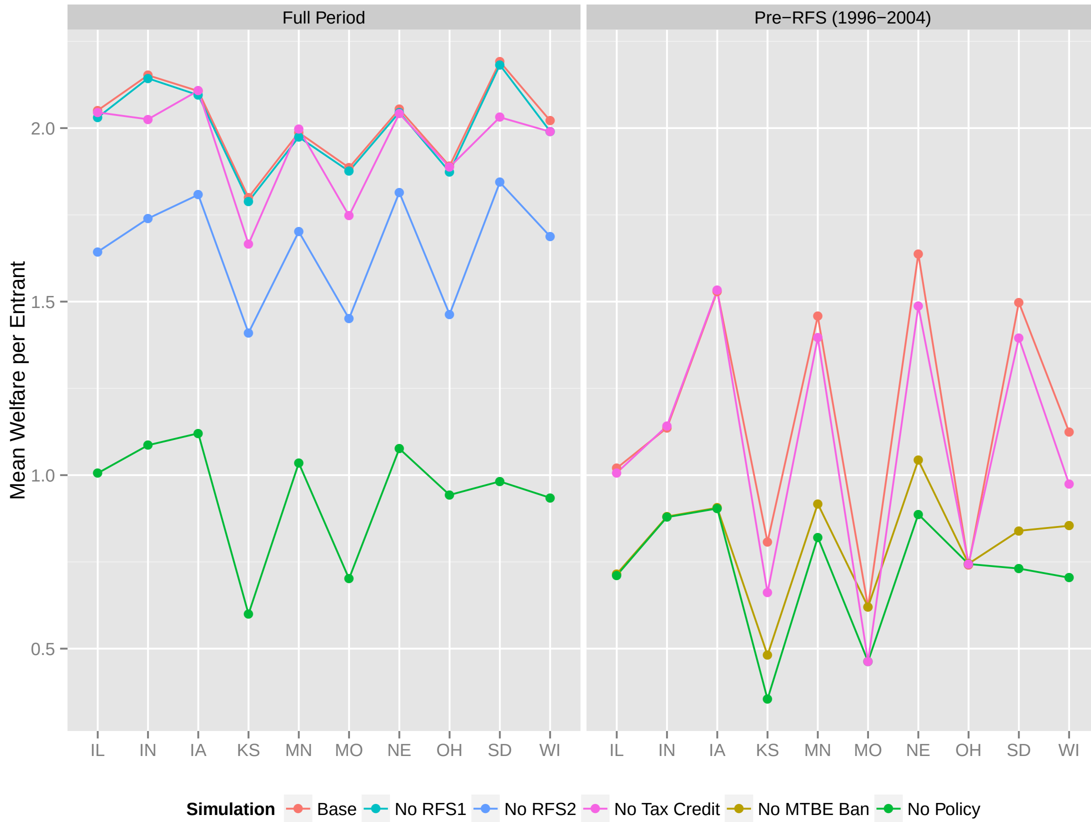

Figure 5: Mean welfare per entrant by state under different policy scenarios