Incomplete Regulation in an Imperfectly Competitive Market: The Impact of the Renewable Fuel Standard on U.S. Oil Refineries

Jesse Burkhardt*
Colorado State University

October 24, 2016

Abstract

I estimate how unexpected increases in the costs of the Renewable Fuel Standard (RFS), a market-based policy, affected U.S. oil refinery prices, markups, marginal costs, and production decisions. The results highlight the risks of failing to account for imperfect competition, multi-product firms, or both when creating policies. First, increases in the costs of the RFS were more than fully passed through to gasoline prices, causing an increase in gasoline markups. Refineries also responded by reallocating production to non-regulated fuels, which led to emissions leakage. Finally, gasoline consumers bore 94% of the burden of the unexpected cost increases.

Keywords: market power; production function; petroleum refining; pass through; emissions leakage.
JEL classification codes: L11, L51, L71, L13, H22, Q48.

---

*Department of Agricultural and Resource Economics, Colorado State University, 1200 Center Ave. Mall, Fort Collins, CO 80523, phone: 503-312-8943, email: jesse.burkhardt@colostate.edu.

Acknowledgments: I am grateful to my committee members Matt Kotchen, Ken Gillingham, and Penny Goldberg for their guidance, patience, and encouragement. I also thank Kerry Smith for being a great mentor and for useful comments. I would also like to thank Jim Bushnell for planting the seeds of this project in my head and Sharat Ganapati for extensive conversations and insights. I want to thank Rahul Deb, Gabe Lade, Meredith Startz, Matthew Grant, Anders Munk-Nielsen, Paige Weber, Stefan Lamp, Nathan Chan, and participants at the Yale Environmental Economics Seminar, NBER Summer Session, and Camp Resources conferences for many useful comments. The data used in this paper were obtained under a researcher information access agreement with the Energy Information Administration, and executed with extensive support from Joseph Conklin of EIA. The views and analysis herein, including any errors, are mine alone.

1 Introduction

When environmental regulations fail to account for existing market distortions, a wide range of unexpected and unintended outcomes can occur (Lipsey and Lancaster 1956). For example, regulations may exacerbate existing welfare losses due to market power by allowing firms to excessively pass the costs of regulation onto consumers (Seade 1985; Weyl and Fabinger 2013). In some cases this may even offset the benefits associated with avoided emissions damages ("External Diseconomies, Corrective Taxes, and Market Structure"). In other contexts, policies can increase production cost inefficiencies across heterogeneous producers (Borenstein, Bushnell, and Wolak 2002). In a similar vein, policies that apply to only a subset of products, referred to as incomplete regulations, allow firms to substitute non-regulated production for regulated production leading to emissions leakage (Fowlie 2009; Auffhammer and Kellogg 2011). The "Theory of the Second Best" – the idea that multiple market failures can interact in ways that make their net effect on welfare different from the effects of each considered in isolation – is a core concept in economics. However, empirical evidence on how this concept plays out in practice is relatively sparse. In this paper, I provide evidence of the effects of incomplete regulation on imperfectly competitive and multi-product firms by evaluating the impact of the Renewable Fuel Standard on the U.S. oil refining industry.

The U.S. Oil refining industry generated over $730 billion in revenue in 2014 and is characterized by high concentration, a complex multi-product production process, and large barriers to entry. By 2014, 50% of the firms that operated refineries in 1986 had exited the market, while at the same time, no new refineries were built and aggregate capacity had increased by 16%. This increase in concentration has generated concern among policy makers that firms enjoy significant and increasing market power. Direct evidence on the extent of market power in the petroleum industry is limited however, as until recently, refinery level cost and pricing information has been unavailable.

The Renewable Fuel Standard (RFS) is one of the most important yet understudied policies currently impacting the oil refining industry. Under the RFS, oil refineries are mandated to blend a certain percentage of biofuels into each gallon of gasoline and diesel sold. Refineries comply by purchasing renewable fuel credits, called RFS credits, from biofuel producers and retiring them with the Environmental Protection Agency (EPA).[1] The current mandates are set to displace 25% of the transportation fuel supply with biofuels by 2022, representing roughly $100 billion in oil company revenue per year (C.F.R. 2015).[2] In doing so, the RFS effectively taxes gasoline and diesel fuels while leaving other petroleum products untaxed.

This paper makes contributions in three areas. First, I modify a recently developed

[1] The RFS credit price is often referred to as the Renewable Identification Number or RIN obligation. In this paper I will refer to it as the RFS credit price.

[2] Statistic for 2012 data. Total sales of wholesale gasoline and diesel was 306,966 thousand gallons per day. Average gasoline prices were $3.68/gallon and average diesel prices were $3.96/gallon. Data retrieved from: http://www.eia.gov/petroleum/data.cfm.

methodology to jointly recover markups, marginal costs, and productivity in the wholesale petroleum product industry.[3] The methodology is unique in that it estimates market power from the firm's cost minimization problem and estimation of a production function. I use this approach for two reasons. Estimating market power is often complicated by a lack of detailed data on marginal costs and well known simultaneity issues. Researchers in industrial organization and international trade have developed a number of strategies to simultaneously estimate markups and marginal costs, which typically rely on structural assumptions about firm behavior and demand conditions. In contrast, the approach I use separately identifies each of these measures and does not require assumptions about consumer demand, market structure, or competition. Instead, the key assumption required to estimate the model is that firms simply minimize production costs.[4]

The richness of my data also allows me to make significant contributions to the petroleum market and production function literature. Researchers do not typically observe plant level data or input allocation across products, which complicates estimating production functions in multi-product settings. An important feature of my data is that I observe the full distribution of physical outputs, which allows me to calculate physical product shares. This detail, combined with the nature of petroleum refining, lets me observe input allocation across end products, which has not been possible in previous work. Moreover, I observe product specific intermediate inputs and the capacities of multiple types of capital for each refinery in the U.S. for more than ten years. These features allow me to directly estimate a multi-product production function and recover productivity at the product level. Furthermore, while others have estimated petroleum refining production functions at the industry level (Berman and Bui 2001), I am the first to estimate a production function at the refinery level.

The second and key contribution of this paper is toward estimating and decomposing the pass-through rate of environmental regulation in regulated and non-regulated petroleum product markets. I find that oil refineries more than fully passed the cost of the RFS onto wholesale gasoline prices but less than fully passed the cost onto ultra-low-sulfur diesel prices in 2013 and 2014. I then use the production function results to decompose the pass-through rate. I find that increases in the RFS credit price increased markups in the gasoline market and marginal costs in the gasoline and diesel markets, indicating that the RFS exacerbated existing market power in the industry. The results are surprising in light of a vast empirical literature that finds less than complete pass-through and constant markups in many contexts (De Loecker et al. 2016; De Loecker and Goldberg 2014; Fabra and Reguant 2014; Goldberg and Hellerstein 2008). In order for firms to over-shift costs, two conditions must be present. First, demand must be log-convex (e.g., constant elasticity). When output decreases in a log-convex demand setting, output prices increase faster than marginal costs. Second, firms must be imperfectly competitive

[3] Markups are the difference between marginal costs and product prices and are a measure of market power.
[4] Additional assumptions are outlined in Section 6

(Seade 1985; Weyl and Fabinger 2013). I estimate mean gasoline and diesel markups of 36% above marginal costs, indicating that firms have significant market power in gasoline and diesel markets. A third condition that likely contributed to excessive pass-through is that in 2013 and 2014, policy uncertainty and technology constraints caused a series of large and unexpected shocks to the RFS credit price (Lade, Lin, and Smith 2015). This indicates that in 2013 and 2014, significant market power combined with unique demand properties allowed firms to benefit from increased credit price volatility by more than fully passing the costs of the RFS onto wholesale gasoline consumers.

An important feature of refineries is that they are multi-product firms by nature. Consequently, regulation can impact non-regulated product prices and markups through at least two channels. Regulations that change the production costs of regulated products will also affect the marginal cost of capital of non-regulated products. For instance, changes in the output mix will have a corresponding effect on non-regulated product marginal costs. Indeed, I find that increases in the RFS credit price caused firms to reallocate production to non-regulated fuels. Specifically, refineries substituted jet fuel production for ultra-low-sulfur diesel production. Consequently, I find that in 2013 and 2014, increases in the RFS credit price caused jet fuel prices and markups to fall, consistent with an outward shift in the jet fuel supply curve. The reallocation of production to non-regulated fuels means some of the avoided emissions gains associated with reduced gasoline and diesel consumption were offset by the corresponding increase in jet fuel consumption. To monetize these losses, I find that a 10% increase in the RFS credit price resulted in an additional \$35-\$179 million in unexpected emissions damages per year. These leaked emissions damages translate to roughly 3-5% of the avoided emissions damages associated with increased biofuel consumption under the RFS.

Third and finally, I use a sufficient statistics approach to estimate the incidence of uncertainty in the RFS credit price. I find that 94% of the burden of the unexpected variation in the RFS credit price in 2013 and 2014 was borne by consumers while only 6% was borne by producers in the gasoline market. This finding can be attributed to the fact that the costs of the policy were more than fully passed through to gasoline product prices during 2013 and 2014. In contrast, I find that only 56% of the burden of the RFS credit price was borne by consumers in the diesel market.

The results of this paper have important policy implications. First, I provide evidence of substantial market power in the petroleum industry, which can be used to inform competition authorities and policy makers generally. Second, the incidence results suggest that shocks in the RFS credit price likely caused substantial short-term producer surplus gains and consumer surplus losses, particularly if retail petroleum distributors fully pass costs onto consumers. Moreover, the long-run goal of the RFS is to encourage innovation in biofuel production and a transition from a non-renewable to a renewable fuel vehicle fleet. These short-run cost shocks suggest that the burden of the transition will fall heavily onto consumers, unless vehicle manufacturers are willing to build more flex-fuel vehicles and fuel providers invest in biofuel filling stations. Third, the results regarding gasoline and diesel markups are especially surprising and important evidence of the interaction between regulation and market power. More broadly, other large polluting industries, such as electricity and concrete, are often considered to be imperfectly competitive and have been shown to exhibit high pass-through rates (Fabra and Reguant 2014; Borenstein and Shepard 2002). This suggests that policy uncertainty may exacerbate market power in other contexts as well, potentially leading to substantial welfare losses due to the sheer magnitude of these industries. Therefore, the benefits of reducing policy uncertainty likely outweigh the investments required to do so. Finally, I show that incompletely regulating multi-product firms can lead to substantial production leakage. The reallocation of production undermines the effectiveness of the policy through at least two channels. Firms adjust their output mix, deviating from cost efficient production, while additional non-regulated production results in emissions leakage. These findings coincide with a large body of literature suggesting that effective policies should simultaneously address emissions and market power externalities.

The remainder of the paper is organized as follows. In the following section, I place my paper in two distinct strands of related literature. Section 3 provides a conceptual framework for the findings in this paper. Section 4 provides some background on the refining process and the Renewable Fuel Standard. In Section 5, I outline the data set I construct and provide some summary statistics. Section 6 describes the methodology to compute markups from the firm's cost minimization problem, while Section 7 outlines the estimation and identification strategies. In Sections 8 and 9 I present results and discussion, and Section 10 provides incidence calculations. Section 11 concludes.

2 Related Literature

In this section I place my paper into two distinct strands of literature: (1) the literature on market power in the petroleum industry and the estimation of production functions, and (2) the literature on energy and environmental policy evaluation.

Market Power and Production Function Literature

The main goal of this paper is to evaluate the impact of the RFS on refinery behavior and to decompose the incidence of the regulation. To do so requires estimation of markups and marginal costs, which I recover using a production function approach. The methodology is based on Hall (1988), and a series of papers in the international trade literature by De Loecker and Warzynski (2012), De Loecker et al. (2016), Collard-Wexler and De Loecker (2015), and De Loecker and Goldberg (2014). In a perfectly competitive market, the elasticity of output with respect to any variable input is equal to that input's share of total revenue. The insight of De Loecker and Warzynski (2012) is that any deviation between these two measures represents a firm's markup over marginal costs. It is this insight which allows me to compute markups from the estimation of a production function. Specifically, I rely only on an estimate of the output elasticity with respect to crude oil inputs.

The richness of my data allows me to immediately address some of the well-known biases associated with estimating production functions. First, I separately observe input and output quantities and crude oil input prices, which alleviates any unobserved price biases. Second, I use the intermediate input demand control function approach developed by Levinsohn and Petrin (2003), based on Olley and Pakes (1996), and recently expanded by Ackerberg, Caves, and Frazer (2015), to address simultaneity and selection issues. Finally, unobserved input allocation presents an additional concern for estimation of multi-product production functions. For example, De Loecker et al. (2016) estimate a production function for single product firms only and develop a routine to recover input allocation for multi-product firms using the single product firm production function estimates. In contrast, I observe the full distribution of outputs from each refinery over time, which allows me to estimate input allocation using physical product shares. I also observe product specific inputs and capacity, which provides additional product level variation needed to directly estimate a product level production function.

Energy and Environmental Policy Evaluation Literature

This paper also contributes to a growing literature on the environmental regulation of the petroleum industry. There is mounting evidence that fuel content regulations under the Clean Air Act resulted in increased prices for regulated fuels. Muehlegger (2006) finds that content regulation contributed to price volatility in California, Illinois, and Wisconsin. Using a difference in difference approach, Brown et al. (2008) find that content regulations are associated with a 3 cents per gallon increase in fuel prices on average. They further relate this increase to market isolation and find that concentration can significantly affect the price differential. Berman and Bui (2001) find that environmental regulations are associated with higher productivity, suggesting that the overall welfare effects of such policies may be understated.

This paper is perhaps most closely related to a recent paper by Sweeney (2015). Sweeney uses a structural model to estimate the effects of the Clean Air Act fuel content regulations on refinery production costs, regulated product prices, and profits. He finds that content regulations increased refinery costs by 7 cents per gallon and 3 cents per gallon for reformulated and low sulfur diesel production, respectively. My analysis differs in that it jointly estimates markups, marginal costs, and productivity without relying on demand side or competitive assumptions. In addition, I estimate markups across the entire distribution of refinery end products, and use the estimated markups and marginal costs to evaluate the incidence of the Renewable Fuel Standard.

One other paper has evaluated the pass-through rate of the RFS credit price to wholesale petroleum product prices. Using non-regulated fuel spot prices as a control for regulated fuel spot prices, Knittel, Meiselman, and Stock (2015) estimate an average long run pass-through rate of 1.01 across diesel and gasoline between 2013 and 2015, with considerable variation at the daily and weekly level. The authors consistently estimate greater than 100% long run pass-through in the diesel spot market and slightly less than complete pass-through in the gasoline spot market. Knittel, Meiselman, and Stock (2015) also estimate short run pass-through and find that 57% of the RFS credit price is passed onto spot prices in the first day rising to 97% pass-through by day 12 on average. There are four main differences between the results presented in this paper and those presented by Knittel, Meiselman, and Stock (2015). First, I estimate pass-through at the rack or bulk distribution terminals, the level at which wholesale transactions occur, while Knittel, Meiselman, and Stock (2015) estimate pass-through in the wholesale spot market. Second, in some specifications Knittel, Meiselman, and Stock (2015) use jet fuel spot prices as a control for aggregate movements in petroleum product prices. I show that non-regulated fuel prices in a multi-product setting are not sufficient controls because firms reallocate production to these fuels. Third, I decompose the pass-through rate using estimates of marginal costs and markups and I use my results to calculate incidence. Finally, I evaluate the impact of the RFS credit price on production decisions.

3 Conceptual Framework

The goal of this paper is to understand how refineries responded to changes in the costs of the RFS. Firms will naturally adjust prices in response to cost shocks but whether or not markups also adjust is an empirical question, the results of which can depend on modeling assumptions. In many contexts, it is common to assume CES demand with monopolistic competition, which leads to constant markups (De Loecker and Goldberg 2014). Under these assumptions, cost shocks such as those experienced in the RFS credit price, will not affect markups. However, these assumptions are unnecessarily restrictive and do not always conform with empirical evidence (De Loecker and Warzynski 2012; De Loecker and Goldberg 2014). Indeed, when firms are imperfectly competitive and demand has sufficient curvature, firms may excessively pass costs onto consumers causing an increase in markups (Weyl and Fabinger 2013; Seade 1985).

To build economic intuition for this possibility and the empirical findings in this paper, consider the markup formula defined by De Loecker and Goldberg (2014),

$$P = \mu(\mathcal{D}, \mathcal{M}) * mc(q, \boldsymbol{z}, \omega, \tau),$$

where P is the price of fuel, $\mu(\mathcal{D}, \mathcal{M})$ is the firm's proportional markup as a function of demand, $\mathcal{D}$, and market structure, $\mathcal{M}$, and $mc(q, \boldsymbol{z}, \omega, \tau)$ is the firm's marginal cost as a function of output, q, production input variables such as input prices, $\boldsymbol{z}$, productivity, ω, and a tax such as the RFS credit price, τ.[5] Shocks in the RFS credit price will directly

[5]Firm and time subscripts are withheld for notational simplicity.

affect marginal costs and may impact productivity in the long-run by incentivizing investment in biofuel blending technology, for example. Likewise, under a variety of demand forms, markups will change in response to cost shocks as a result of incomplete or more than complete pass-through.

Figure 1 shows how this effect might occur with an increase in the RFS credit price, or tax rate τ, resulting in an upward shift in marginal costs.[6] The graph on the left illustrates pass-through with linear demand while the graph on the right illustrates pass-through with log-convex demand. In these two stylized graphs, pass-through is the change in price, ΔP, relative to the change in marginal costs, ΔMC, given a change in the tax, τ. With linear demand (the graph on the left), the before tax markup, μ, is clearly larger than the after tax markup, μ_τ, which is a function of marginal cost increasing more than price in response to the tax, or less than complete pass-through. In contrast, when demand has some curvature (the graph on the right), the after tax markup, μ_τ, is clearly larger than the before tax markup, μ, which is a function of price increasing more than marginal cost in response to the tax, or more than complete pass-through.

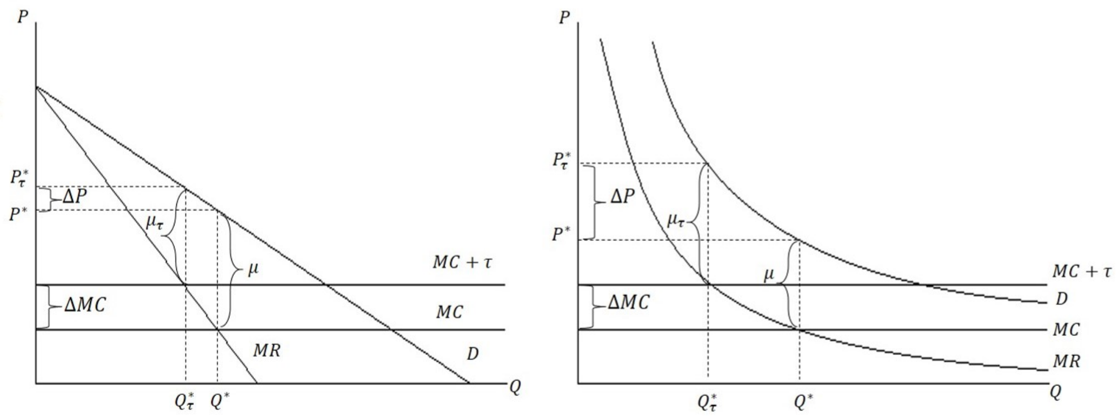

Figure 1: Monopoly Pass-Through with Linear Demand and Demand with Curvature

Given these graphical results, the remaining question is the following: given a change in tax rate, τ, under what conditions will pass-through be greater than one, $\frac{dP}{d\tau} > 1$. In general, Weyl and Fabinger (2013) show that monopoly and oligopoly pass-through depend not only on supply and demand elasticities, but more importantly on the curvature of demand.[7] The intuition is that imperfectly competitive firms set marginal revenue equal to marginal costs where marginal revenue is a function of demand. If demand is log-convex (e.g., constant elasticity), the slope of the demand curve, and therefore the slope of the marginal revenue curve, is increasing in price. When this occurs, pass-through

[6]For illustrative purposes, marginal costs are assumed to be constant.

[7]For example, Weyl and Fabinger (2013) show that monopoly pass-through can be written $\frac{dP}{d\tau} = \frac{1}{1+\frac{\epsilon_d - 1}{\epsilon_s} + \frac{1}{\epsilon_{ms}}}$, where ϵ_s is the elasticity of supply and ϵ_{ms} is the elasticity of marginal consumer surplus or the firm's absolute markup, $ms = -P'q$. Weyl and Fabinger (2013) show the formula can be generalized to any form of competition. Similarly, Seade (1985) shows that excessive pass-through can occur in an oligopoly setting if firms compete in quantities and the elasticity of the slope of the demand curve is greater than one. For example, when demand is isoelastic, the elasticity of the slope of the demand curve is $E = 1 + 1/\epsilon_d$.

will be greater than one causing markups to increase with reductions in output (Seade 1985; Weyl and Fabinger 2013). These results generalize to other forms of competition.

4 Industry and Policy Background

In the following section, I provide a brief overview of the refining industry, the refining process, and the Renewable Fuel Standard.

U.S. Oil Refining

There are 155 refineries in the U.S. owned by 63 corporations in 31 states (EIA 2014a). The regions with the largest concentration of refineries are along the Gulf Coast, which includes Texas (27 refineries) and Louisiana (19 refineries), and along the West Coast, which includes California (18 refineries) (EIA 2014a). There have been no new large refineries (greater than 100,000 barrel per day capacity) opened in the U.S. since 1977 (EIA 2014b). At the same time, Figure 2 shows that a substantial number of refineries have exited the market since 1986 while the remaining refineries have increased capacity, resulting in a dramatic increase in industry concentration. The number of refineries exiting and the number of mergers stabilized in the early 2000's, which alleviates concerns about entry and exit selection bias. These facts, combined with geographic market isolation (markets are isolated due to geographic boundaries and pipeline constraints), the distribution of small to large refineries, and evidence from the economics literature has led to widely held beliefs that the petroleum industry is imperfectly competitive (Muehlegger 2006; Sweeney 2015; Borenstein and Shepard 2002).

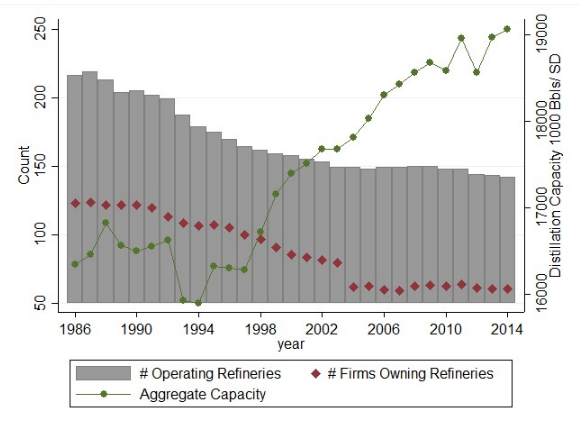

Figure 2: Number of Refineries, Firms Owning Refineries, and Aggregate Capacities

To estimate a refinery production function, one first needs to understand the production process. The first step of the petroleum product supply chain involves refineries purchasing a mix of imported and domestic crude oil. Crude oil is essentially a mixture of heavier and lighter hydrocarbons. The two primary components that define crude quality are weight (specific gravity or API - the ratio of light to heavy hydrocarbons) and sulfur content. Low specific gravity (light) and low sulfur (sweet) crude oils are higher quality and require less effort to refine, but are more expensive to purchase. Heavy, inexpensive crude oil requires more processing for the equivalent amount of end product.

Once the crude oil is purchased, the distillation process begins. Input quantity adjustments are costly so each refinery makes production runs that are specifically calibrated for a given type of crude oil and a desired mix of end products.[8] In each production run, refineries choose a mix of end products to maximize current profits based on exogenous factors such as crude input quality and prices, forecasted demand for each product, and endogenous factors such as the costs of refining. A refinery can choose to exert more effort towards producing a given end product, conditional on input quantity and existing capacity, but at increasing costs. For example, conditional on existing capital and crude oil quality, refineries can adjust the end product mix by approximately 2-6%.[9] On the other hand, refineries can change their output mix more substantially by investing in more costly capital. Data on the downstream capacity of each refinery in the U.S. provides additional product level variation to identify the parameters of the production function.

All of the crude oil entering the refinery is first processed in the atmospheric distillation tower. The distillation tower heats the crude oil allowing heavier and lighter hydrocarbon chains to naturally separate. Lighter hydrocarbons, like those immediately suitable for gasoline, naturally rise to the top or evaporate and are siphoned off. Heavier hydrocarbons are removed from various heights along the side of the distillation unit. The heavier hydrocarbons can be further treated in downstream units, such as the catalytic cracker, to be converted into lighter, more profitable hydrocarbons at additional costs.[10]

Refineries convert virtually all of the hydrocarbons in a barrel of crude oil into end products ranging from the lowest residual fuel oils to diesel, gasoline, and liquified refinery gases. For example, catalytic reforming dehydrogenates heavier fuels creating excess hydrogen gas, which is then used in other refining processes such as hydrocracking (Gary, Handwerk, and Kaiser 2007). This is an important point for estimation because it implies that the total volume of end product output is very closely related to the total volume of crude input. However, the input to output ratio is not unity because, 1) there are a handful of additional additives that are blended with fuels such as oxygenates, which are used to upgrade gasoline, and 2) petroleum expands in volume during the refining process.

Refineries report an atmospheric crude distillation capacity at the beginning of each month. This measure is based on the profit-maximizing production run for a given

[8]Production runs typically last 2-4 weeks (Gary, Handwerk, and Kaiser 2007).

[9]In a phone call with a Chevron refinery in California, I was told that refineries could alter their production ratios by about 2 percentage points without substantial investment in new infrastructure. Gary, Handwerk, and Kaiser (2007) reports 6%.

[10]The heavier fuels can be processed in many different types of processing units including a catalytic reformer, fluid catalytic cracker, hydrocracker, and delayed coker. Each unit works in a slightly different way. For example, naphtha is treated in the catalytic hydrodesulfurizer to remove excess sulfur and then in a catalytic reformer to “reform” naphtha molecules into more complex molecules with higher octane ratings. Octane simply refers to the number of carbon atoms in a hydrogen molecule.

refinery and for a specific type of crude, meaning the self reported capacity measures are potentially correlated with output choices. Changes to capacity are rare and only occur when substantial upgrades or downgrades are made to a refinery.[11]

Policy Background: The Renewable Fuel Standard

The Energy Policy Act of 2005 established the RFS under the umbrella of the Clean Air Act. The policy seeks to increase domestic biofuel consumption to 36 billion gallons (bgals) per year by 2022 by mandating that the total volume of gasoline and diesel sold in the U.S. is blended with a minimum volume of renewable fuel. The blending proportion is set annually by the EPA and is referred to as the blend mandate. Additional goals of the RFS are to significantly reduce greenhouse gas emissions from the consumption of transportation fuels, and to increase energy security by reducing petroleum imports (EPA 2015a). This section outlines how the RFS credit price is calculated and why variation in the credit price can be thought of as exogenous to refineries.

The EPA keeps track of the quantity of renewable fuel blended with conventional fuel via a system of tradable credits. Each gallon of renewable fuel that is produced in the U.S. or imported to the U.S. generates a renewable fuel credit, called a Renewable Identification Number (RIN). Obligated parties under the RFS (petroleum refineries, petroleum importers, and blenders) purchase RINs from renewable fuel producers. The RIN is detached from the renewable fuel when the renewable fuel is blended with conventional fuel. The obligated parties must retire RINs to the EPA in proportion to the quantity of conventional gasoline and diesel that they produce. If the obligated party has a surplus of RINs they can sell excess RINs to other obligated parties that have a deficit, creating a market for RINs. Thus, RIN trading is a transfer payment between refineries and biofuel producers and effectively taxes gasoline and diesel production while subsidizing biofuel production. It is important to note that only gasoline and diesel are regulated under the RFS, while other products such as jet and aviation fuel are unregulated.

The RFS specifies four nested categories for renewable fuels: total or conventional renewable fuels (such as ethanol), advanced biofuel, biomass-based diesel (BBD), and cellulosic.[12] Each of the four categories is associated with a category specific RIN and a category specific blending requirement, or blending percentage. However, the nested structure of the blending mandate allows cellulosic and biodiesel RINs to count towards the advanced biofuel mandate, and advanced biofuel (and biodiesel and cellulosic) RINs to count towards the total biofuel mandate. The EPA calculates these fractions based on the desired biofuel consumption in a given year, divided by the total projected domestic

[11] A note on capacity measures is warranted. In industries such as oil refining, capacity is not measured homogeneously across firms but is based on an optimal product mix (Cowing and Smith 1977). This measurement error will produce inconsistent capacity coefficient estimates but will not affect the coefficients of interest if input choice is uncorrelated with capacity measurement.

[12] Cellulosic fuels are biofuels produced from non-edible portions of plants, biodiesel is commonly produced from soybean or canola oil, advanced biodiesel is biofuel with life-cycle emissions at least 50% below baseline values, and the overall renewable biofuel is all approved biofuel including biofuel produced from cornstarch such as ethanol.

transportation fuel consumption in that year. Importantly, the blending mandate does not incentivize firms to produce more gasoline or diesel because firms can substitute credits across fuels.

The price of the four RINs can be aggregated to an overall RIN price obligation, which I will refer to as the RFS credit price. For example, in 2013, the blending standards required that for each gallon of gasoline or diesel sold, 0.0005 cellulosic RINs, 0.0113 biomass-based diesel (BBD) RINs, 0.0162 advanced RINs, and 0.0974 conventional renewable fuel RINs were to be retired (C.F.R. 2015). Note that because of the nested structure of the policy, the biodiesel mandate counts towards the advanced and total biofuel mandates so that the total 2013 biofuel mandate can be met by turning in $0.0812 = 0.0974 - 0.0162$ RINs, for example. Therefore, the aggregate 2013 and 2014 RFS credit price per gallon of gasoline or diesel sold by an obligated party is:

$$P_{RIN}^{2013-14} = 0.0113 P_{RIN}^{BBD} + 0.0049 P_{RIN}^{Adv} + 0.0812 P_{RIN}^{RFS}$$

where P_{RIN}^{BBD}, P_{RIN}^{Adv}, and P_{RIN}^{RFS} are the prices of the BBD, advanced, and conventional biofuel RINs respectively.[13][14]

A key identifying assumption in this paper is that shocks to the RFS credit price in 2013 and 2014 were exogenous to refinery decisions. Figure 3 shows the aggregate RFS credit price for 2011 through 2014. Prior to 2013, the RFS credit price was low and fairly stable. However in 2013 there was a substantial spike in the RFS credit price, with some volatility carried through to 2014. To understand the shock in the credit price in 2013, it is important to understand the nature of ethanol blending. Ethanol is blended with gasoline at three main levels: E0 containing 0% ethanol; E10 containing 10% ethanol; and E85 containing roughly 70-85% ethanol. The vast majority of vehicles on the road today can burn fuel that contains up to 10% ethanol. Going beyond the 10% level to 11% or 12% ethanol can damage existing engines. The limit of 10% ethanol is commonly referred to as the blend wall and is the primary reason for the shock in the credit price in 2013.

When the RFS blending mandate is below the blend wall, as it was prior to 2013, a nonzero credit price can be attributed entirely to transaction costs (Burkholder 2015). The reason is that the RFS credits are a subsidy payment to the ethanol producers and are equal to the difference between the supply price of ethanol and the demand price for ethanol. As such, the credit price, and therefore the subsidy payments, largely depend on the marginal gallon of ethanol sales (Burkholder 2015). When the RFS mandate is below the blend wall, the marginal gallon of ethanol is sold as E10. E10 contains 3% less energy per volume than E0 so refineries and blenders can sell E10 at virtually the same

[13] In practice, the formula is adjusted for the blending mandates of different years. In 2011, the mandates were 0.0069, 0.0078, and 0.081 and in 2012, the mandates were 0.0091, 0.0121, and 0.0923 for BBD, advanced, and conventional renewable fuels respectively (C.F.R. 2015).

[14] As is common in the literature, I ignore the cellulosic mandate (Knittel, Meiselman, and Stock 2015; Lade, Lin, and Smith 2015). The blending requirement for cellulosic fuels is much lower than the other requirements meaning a minor amount of the renewable fuels blended into the market have been cellulosic fuels.

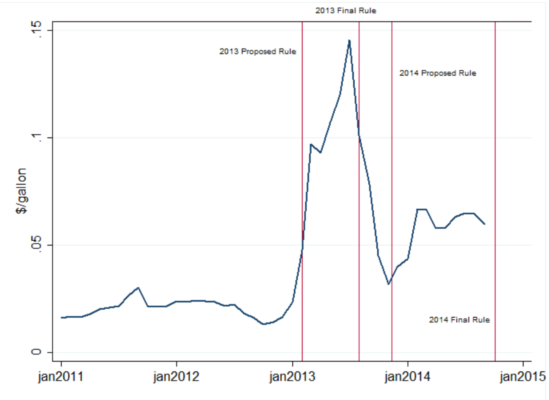

Figure 3: RFS credit price per gallon of gasoline and diesel produced

price as E0. However, if the blend wall is breached, as was initially proposed by the EPA in early 2013 (1st solid red line in Figure 3), the marginal gallon of ethanol must be sold as E85. E85 contains 33% less energy per volume than E0 meaning a car burning E85 will travel a noticeably shorter distance than a car burning E0. Consumers are therefore willing to pay less for a gallon of E85 than a gallon of E0 or E10. Thus, the demand price for ethanol is relatively high when the marginal gallon of ethanol sales is E10 and relatively low when the marginal gallon of ethanol sales is E85. This implies the subsidy payment, and therefore the RFS credit price, increases when the blend wall is breached and E85 is the marginal fuel. Additionally, ethanol and biodiesel RIN prices converged in 2013, which suggests that refineries responded by over complying with the biodiesel mandate as some biodiesel RINs can be substituted for ethanol RINs (Lade, Lin, and Smith 2015; Irwin 2014).

The fact that the 2013 proposed rule was expected to breach the blend wall explains the initial increase in the credit price in 2013 but does not fully explain the decrease in the credit price in the latter half of 2013 or the subsequent variation in 2014. As discussed in Knittel, Meiselman, and Stock (2015) and Lade, Lin, and Smith (2015), the additional volatility in the RFS credit price was brought on by policy uncertainty. In particular, Lade, Lin, and Smith (2015) shows that the largest drivers of the variation in RFS credit prices were three separate policy shocks: the release of the EPA's 2013 Final Rule (2nd vertical line in Figure 3), which caused the decrease in the credit price in the latter half of 2013; a leaked version of the EPA's 2014 Proposed Rule, which caused a further decrease in the credit price in late 2013; and the release of the 2014 Proposed Rule (3rd vertical line in Figure 3), which caused the credit price to increase once again.

This evidence suggests that the variation in the RFS credit price in 2013 and 2014 was caused by policy uncertainty regarding whether or not the EPA's blending mandates would breach the blend wall, which is a technology constraint in the vehicle fleet. The credit price variation is therefore exogenous to refineries because the refineries have no short term control over the composition of vehicles on the road.

5 Data

I construct a confidential refinery-firm level data set spanning 2004-2014 using surveys from the U.S. Energy Information Administration.[15] Production data is collected at the refinery level while sales data, including input and output prices, are collected at the firm-region level. For example, firms such as Chevron, may own multiple refineries in different locations around the U.S.

Survey form EIA-810 provides very detailed data on each refinery's inputs, gross production, gains and losses, shipments, ending stocks, and capital. I observe the crude oil inputs in thousands of barrels, as well as the average API gravity and sulfur content of the crude oil used in a given month. In addition, I observe the full distribution of products produced by each refinery in each month. This allows me to construct a measure of the share of inputs allocated to the production of each output, the importance of which will be discussed in the following section. Each month, refineries report the output of approximately sixty end products, most of which fall into several broad categories including liquified petroleum gases, aviation fuel, gasoline and gasoline blending components, jet fuel and kerosene, distillates (diesel fuels), heavy residual fuel oils, asphalt and road oil. Gasoline and diesel are reported by various types including conventional and reformulated gasoline and high, low, and ultra-low-sulfur diesel. Refineries also report the inputs of each petroleum product and blending components such as oxygenates, biofuels, or unfinished oils. I subtract petroleum product inputs such as unfinished gasoline, diesel, and kerosene from the gross production of finished gasoline to construct net production of finished fuels.[16]

At the annual level, survey form EIA-820 provides information on the capacity of the distillation tower and select downstream processing units. I observe the downstream fractioning capacity of the vacuum distillation unit and four thermal cracking units (including two coking units): the catalytic cracking unit, and three hydrocracking units separated by residual, distillate, and gas oil cracking capacity. I also observe the downstream capacity of the reformer, the capacities of the heavy gas oil and naphtha hydrotreaters, and separate desulfurization capacities for gasoline, kerosene and jet fuel, diesel, other distillates, residual fuel oils, and all other fuels.

At the firm-Petroleum Administration for Defense District (PADD) level, I observe crude oil input prices for domestic and imported crude from survey form EIA-14. Firms report sales prices of gasoline, diesel, jet fuel, aviation fuel, and a handful of other products by state, fuel type (regular, mid-grade, premium), and sales type (retail, rack, dealer-tank-wagon, bulk, commercial/industrial, and other end users) on form EIA-782A.[17]

[15] Surveys can be found at http://www.eia.gov/survey/.

[16] This is how the EIA estimates net production of fuels. See the definition of refinery production here: http://www.eia.gov/dnav/pet/tbldefs/pet_pnp_refp2_tbldef2.asp

[17] Rack prices are the prices paid at the terminal for deliveries of end product in truckload sized quantities. Dealer-tank-wagon prices are essentially forward contract prices. The dealer-tank-wagon prices are consistently higher than other prices due to the guarantee of sales, regardless of supply disruptions. Bulk prices are assigned to bulk sales larger than a truckload. All sales on form EIA-782A are reported in the state where the transfer of title occurred. The transfer of title

Petroleum products travel around the U.S. via pipeline, tanker, and truck. The majority of petroleum products leave refineries via pipelines and make an intermediate stop at a terminal (bulk storage facility) where they are temporarily stored, blended with biofuels, and then trucked to retail gasoline and diesel stations, or commercial customers.[18] The sales prices reported on EIA-782A do not include taxes but do include shipping costs. I follow Sweeney (2015) to construct an estimate of shipping costs for each firm. Firms are assumed to minimize transportation costs by supplying each state with end product produced from the nearest refinery. I use a GIS mapping tool to find the distance between each refinery and each terminal in each state following pipelines. I assign a transportation cost of 2 cents per gallon per thousand miles traveled (Sweeney 2015; Muehlegger 2006). I can then use these estimates to subtract transport costs from the sales prices reported on EIA-782A.

The refinery and firm level data from the EIA is supplemented with data from the U.S. Bureau of Labor Statistics. I construct monthly-state level labor use by multiplying annual state level total employment by trends in monthly national level employment for petroleum refinery operators. Specifically, I use the total number of employees multiplied by the average number of hours worked per week.[19]

Summary Statistics

Table 1 presents refinery level production data summary statistics. I observe production data for 155 refineries owned by 65 firms between 2004 and 2014. The largest refinery in the U.S. can process over 600,000 barrels of oil per day while the smallest refinery can only process 33. Hence, refineries are quite heterogeneous and the distribution of capacity suggests the industry is composed of a subset of strategic producers with a competitive fringe. The average and median net production to crude oil input ratio is 1.25 and 1.1 respectively, which indicates that the median production output is 10% greater than crude oil inputs. This could be due to expansion in the refining process, inputs of unaccounted for unfinished oils, and inputs of fuel additives.[20] Downstream capacities show an intuitive trend. On average, gasoline has the largest downstream capacity followed by diesel, jet fuel, and other fuel.[21] The last three rows show the average inputs of oxygenates, which are used to upgrade gasoline, average inputs of renewable fuels, which are blended with gasoline and diesel, and average kerosene and unfinished kerosene type oils used to create jet fuel. The table also presents average inputs and outputs, average crude quality, and average state level labor of petroleum refinery operators. Average crude oil input prices

typically takes place at distribution terminals but the end product could ultimately be consumed in a neighboring state. I use rack, bulk, and sales for resale prices.

[18] There are over 192,000 miles of pipelines in the U.S. In 2013, 96% of all products sold were shipped via pipeline (in total, 6.6 billion barrels of natural gas and petroleum products). Source: http://www.eia.gov/dnav/pet/pet_cons_psup_dc_nus_mbbl_a.htm and http://www.aopl.org/wp-content/uploads/2014/10/U.S.-Liquids-Pipeline-Usage-Mileage-Report-Oct-2014-s.png

[19] Implications of the labor variable in estimation are discussed in Section 7

[20] On average, petroleum expands about 6% by volume during the refining process.

[21] Note that some small refineries do not have downstream processing capacity.

are presented in Table 9. Interestingly, the West Coast and Gulf Coasts have the highest reported crude oil input prices.

Table 1: Production Data Summary Statistics (2004-2014)

Variable	Mean	Std. Dev.	Min.	Max.	N
Output to Input Ratio	1.249	7.5	0.762	913.5	90,265
Crude Inputs (1000's of bbls)	4103.445	3433.527	0	18495	90,265
Net Production (1000's of bbls)	4783.703	4021.672	0	21576	90,265
Crude Oil Price ($/bbl)	77.192	24.793	24.913	137.695	88,020
API Gravity	31.593	7.427	9.9	54.6	90,265
Sulfur Percent	1.259	0.936	0.01	7.03	90,265
Total Employment (Employees/State)	4759.193	4902.097	29.181	14484.452	90,265
Atmospheric Capacity (bbls/CD)	150441.467	120916.009	33	600250	90,265
Downstream Gas Cap (bbls/SD)	106894.684	101305.883	0	3148213.75	90,265
Downstream Diesel Cap (bbls/SD)	80891.791	80655.283	0	3180712.75	90,265
Downstream Jet Cap (bbls/SD)	58718.442	70642.146	0	3095615	90,265
Downstream Other Cap (bbls/SD)	37869.02	61485.578	0	1797066.375	90,265
Oxygenates Inputs (1000's of bbls)	251.774	437.653	0	3071	90,265
Renewable Fuel Inputs (1000's of bbls)	6.361	19.658	0	279	90,265
Kerosene UFO Inputs (1000's of bbls)	67.072	172.146	0	1634	90,265

Notes: bbls/CD represents barrels per calendar day and bbls/SD represents barrels per stream day. In some cases firms only report one or the other measure. Zero inputs or outputs can be attributed to refineries reporting zero inputs or outputs for a given month, possibly due to scheduled or unscheduled shutdowns. Not all refineries have all downstream processing machines so in some cases downstream capacity is zero. The number of observations for crude oil prices is less than the number of observations for the other variables because the crude oil price is observed at the firm-PADD level whereas all other variables are observed at the refinery-product level. However, the crude oil price is not used in the production function estimation. It is only used to compute markups.

Tables 2 and 3 present end product price summary statistics by region and fuel type. The prices represent average wholesale prices (rack, bulk, or sales for resale prices). Not surprisingly, average prices are highest along the East Coast (PADD 1) and West Coast (PADD 5) and lowest in the interior of the U.S. and along the Gulf Coast (PADD 3). The West Coast has the some of the strictest fuel content regulations while the Gulf Coast region has the greatest number of refineries. Consistent with previous research, reformulated gasoline prices are higher than conventional gasoline prices on average. Jet fuel is priced similar to reformulated gasoline while low end fuels have the lowest price on average.

Table 2: Output Price Summary Statistics by Region in $/gal

PADD	Mean	Std. Dev.	Min.	Max.	N
East Coast (1)	2.177	0.696	0.536	4.062	3,671
Midwest (2)	2.111	0.755	0.501	4.130	8,854
Gulf Coast (3)	2.115	0.736	0.418	4.104	8,982
Rocky Mountain (4)	2.033	0.893	0.460	4.249	4,693
West Coast (5)	2.240	0.739	0.479	4.435	6,305
Total	2.133	0.765	0.418	4.435	32,505

Table 3: Output Price Summary Statistics by Fuel in $/gal

Fuel	Mean	Std. Dev.	Min.	Max.	N
Conventional Gasoline	2.286	0.638	0.855	3.992	6,371
Reformulated Gasoline	2.346	0.634	0.918	3.839	3,863
Mid and Low Sulfur Diesel	2.196	0.769	0.868	4.249	3,181
Ultra Low Sulfur Diesel	2.572	0.631	0.897	4.435	5,269
Jet Fuel	2.394	0.700	0.866	4.148	4,016
Low End	1.587	0.667	0.418	4.077	9,805
Total	2.133	0.765	0.418	4.435	32,505

6 Framework to Estimate Firm-Product Level Markups

The main goal of this paper is to understand the impact of unexpected shocks in the RFS credit price on oil refinery markups, marginal costs, product prices, and production decisions. Doing so requires the explicit estimation of markups and marginal costs. Often in the literature, estimation of market power depends on structural assumptions about the shape of the demand curve, the nature of competition, and market structure. Instead, I modify a novel approach originally developed by Hall (1988), and more recently expanded on by De Loecker and Warzynski (2012), De Loecker (2011), and De Loecker et al. (2016), that relies on two simple key assumptions: that firms minimize production costs and that input allocations are observed.

To obtain markup estimates, the approach relates the output elasticity of a variable input with that input's share of expenditures in total sales. Accordingly, I need to estimate a production function to recover output elasticities. In the following section, I first show how markups are derived from production data and outline the assumptions I make to do so. I then discuss the estimation and identification of output elasticities. Finally, I show how markups and marginal costs are computed from estimated output elasticities and data on input expenditures and total sales.

Derivation of Markups

Assume a cost minimizing firm f producing product j faces the following Lagrangian summed over all refineries i owned by firm f at time t,

$$\mathcal{L}_f(\boldsymbol{V}_{ijt}, \boldsymbol{K}_{ijt}, \lambda_{ijt}) = \sum_{i \in I_f} \left(\sum_{v=1}^{V} W_{ft}^{v} V_{ijt}^{v} + \sum_{k=1}^{K} W_{ft}^{k} K_{ijt}^{k} + \lambda_{ijt} \left[Q_{ijt} - Q_{ijt} \left(\boldsymbol{V}_{ijt}, \boldsymbol{K}_{ijt}, \Omega_{it} \right) \right] \right),$$

where Q_{ijt} is net physical output of product j produced by refinery i, which is owned by firm f at time t, $\boldsymbol{V}_{ijt}$ is a vector of variable inputs including crude oil, C_{ijt}, labor, L_{ijt}, and fuel additives, and $\boldsymbol{K}_{ijt}$ is a vector of dynamic inputs such as capital.[22] Note that

[22]To understand net physical output, see section 5. The setup and assumptions closely follow those in De Loecker et al. (2016) but are tailored to fit the present setting.

variables are indexed by either firm f or refinery i depending on the level of observation in the data. Let the price of variable inputs for firm f owning refinery i be denoted W^v_{ft} for $v = \{1, ..., V\}$, and similarly let W^k_{ft} denote the price of dynamic inputs for $k = \{1, ..., K\}$. The first order condition for the crude oil used to produce product j at refinery i is

$$W^c_{ft} = \lambda_{ijt} \frac{\partial Q_{ijt}(\cdot)}{\partial C_{ijt}}.$$

The marginal cost of producing product j at refinery i, conditional on observed output Q_{ijt} is λ_{ijt}. Rearranging and multiplying both sides by $\frac{C_{ijt}P_{fjt}}{Q_{ijt}}$ yields

$$P_{fjt}\left(\frac{\partial Q_{ijt}(\cdot)}{\partial C_{ijt}}\frac{C_{ijt}}{Q_{ijt}}\right) = \frac{P_{fjt}}{\lambda_{ijt}}\frac{W^c_{ft}C_{ijt}}{Q_{ijt}}, \tag{1}$$

where P_{fjt} is the output price for product j produced by refinery i owned by firm f. The term in parentheses on the left hand side of (1) is the elasticity of output for product j with respect to the share of crude oil input, C_{ijt}, allocated to producing product j, and will be denoted θ^c_{fjt}.

Define firm-product-time level markups as the ratio of output prices and marginal costs, $\mu_{fjt} = \frac{P_{fjt}}{\lambda_{fjt}}$. Rearranging (1) yields an expression for firm-product-time specific markups as a function of the output elasticity, θ^c_{fjt}, and the ratio of the revenue associated with product j to the share of input expenditure devoted to producing product j

$$\mu_{fjt} = \theta^c_{fjt}\left(\frac{P_{fjt}Q_{fjt}}{W^c_{ft}C_{ijt}}\right). \tag{2}$$

The expression in (2) is analogous to one derived by De Loecker et al. (2016).[23] Although this markup derivation is now common in the literature, my contribution is to estimate refinery-product specific elasticities allowing me to compute a rich set of firm-product level markups. In the present context, both terms on the right hand side of (2) are unobserved and must be estimated from the data.

Model Assumptions

Two key assumptions are used to operationalize the model. First, I assume that refineries minimize short-run costs conditional on observed profit maximizing output. This implies output choices are an economic decision while input allocation is an engineering decision. The conditionality of this assumption also bridges the duality gap between profit maximization and cost minimization for imperfectly competitive firms.

Cost minimization requires that refineries are price takers in the input market. It is widely believed that the crude oil market is highly competitive, at least from the buyers

[23] All variables are indexed by either firm, f, or refinery, i, depending on the level of observation in the data. For notational simplicity, I omit regional (PADD) level subscripts. However in practice, I assign firm $f's$ input prices for PADD X to firm $f's$ refineries in PADD X. Therefore, in any equation with firm and refinery level variables, a representative refinery, i , is owned by a representative firm, f.

side. To explain observed variation in input prices, I assume refineries face a vector of crude oil prices observed at the firm-PADD level $W_{ft}^c = W_t^c(\boldsymbol{\nu}_{it}, D_{it}, G_i, a_{fjt-1})$, where prices depend on a vector of crude quality measures $\boldsymbol{\nu}_{it}$, for instance API gravity and sulfur content, the origin of the crude oil D_{it}, i.e., domestic or international, the refinery's location, G_i, and firm level actions taken in periods prior to period t. These actions could encompass pre-negotiated input price contracts so long as the contracts do not specify prices as a function of input quantity (De Loecker et al. 2016).

The second primary assumption addresses unobserved input allocation. I assume that the share of observed variable and fixed inputs are attributable to observed outputs. This assumption allows me to use physical output shares as an estimate of input shares and applies to both the production function estimation and the computation of markups.[24] In most production contexts, assigning inputs to outputs is virtually impossible. In contrast, the petroleum production process is uniquely transparent. A barrel of crude oil contains a certain quantity of molecules composed of carbon atoms, hydrogen atoms, oxygen atoms, sulfur atoms, and trace amounts of other elements. The process of refining crude oil rearranges these atoms to form different, more profitable molecules. Although some expansion occurs in the refining process, and additional fuel additives and blending components are added to final products, the crude oil input to total output ratio is close to one.[25] Therefore, the total volume of liquid entering a refinery is roughly equivalent to the total volume of liquid exiting the refinery.[26] For this reason, the physical output ratio is a good approximation of the quantity of crude oil allocated to produce each end product. I thus estimate the share of inputs allocated to the production of product j as

$$\rho_{ijt} = \frac{Q_{ijt}}{\sum_j Q_{ijt}}. \tag{3}$$

I use atmospheric distillation capacity as the main measure of a refinery's capacity. The proportion of the atmospheric distillation tower devoted to producing a given end product can also be approximated using the observed physical product shares because all of the crude oil entering a refinery passes through the atmospheric distillation tower. For the downstream processing units, I assign product shares based on the products produced by those machines. For example, the reforming units only produce gasoline so the reforming capacity is multiplied by the within gasoline product shares (i.e., reformulated and conventional gasoline). Further implications of this assumption will be addressed in

[24] There are three ways to address unobserved input allocation: eliminate multi-product firms from the dataset, aggregate production to the firm level, or assume a method for allocating inputs across products (De Loecker and Goldberg 2014). The first is infeasible in the refining context as no refinery produces only a single product. The second results in a loss of efficiency and the ability to estimate product specific production functions. I follow the third method, which has been employed by a number of researchers. Foster, Haltiwanger, and Syverson (2008) allocates input expenditures according to revenue shares and De Loecker (2011) allocates them based on the number of products. De Loecker et al. (2016) develop a method to recover multi-product firm product shares using single product firms, resulting in allocations that are similar to the physical input allocations used in this paper. Ultimately, many authors aggregate production to the firm level because assigning inputs to outputs is nearly impossible in many industries (De Loecker 2011).

[25] The median input to output ratio in my data is 1.1. Other molecular losses might come in the form of emissions output, i.e., CO_2.

[26] For instance, a refinery cannot use 30% of total crude oil inputs to produce 90% of outputs.

greater detail in the following section.

Equation (3) also ensures that the input allocations sum to 1. This constraint captures the multiproduct aspect of the oil refining production process. If a firm allocates more crude oil or capacity to one product, it must allocate less crude or capacity to a different product. In doing so, the firm changes the total production of both products, which in turn affects the marginal costs of both products. In other words, I estimate marginal costs, and therefore markups, at a particular point in the production process rather than the marginal cost curve. The benefit of this methodology is that it is completely flexible and allows marginal costs to be decreasing, increasing, or constant in quantity and captures the multi-product nature of production in a simple framework.

Finally, estimating a multi-product production function requires two additional assumptions. As is common in the literature, I assume the production technology is common across the set of producers and the production function is continuous and twice differentiable with respect to crude oil inputs. The former assumption implies all refineries producing product j do so using the same technology, conditional on the type of technology (downstream processing machine) being used. For example, each refinery has an atmospheric distillation unit but not all refineries have each type of downstream processing unit. The latter assumption rules out fixed proportion or Leontief technology for crude oil. Indeed, firms have some degree of flexibility in the end product mix from a barrel of crude oil. Importantly, none of the above assumptions imply output elasticities are constant across firms or products, except in the special case of Cobb-Douglas.

7 Obtaining Output Elasticities: Production Function Estimation and Identification

Given the assumptions outlined above, the general expression for the refinery level production function is the following

$$Q_{ijt} = F\left(C_{ijt}, K_{ijt}, DK_{ijt}, L_{sjt}, \boldsymbol{Z}_{ijt}; \boldsymbol{\beta}\right)\Omega_{ijt}, \tag{4}$$

where C_{ijt} is the crude oil input used by refinery i to produce product j at time t, K_{ijt} is the atmospheric distillation capacity of refinery i producing product j at time t, DK_{ijt} is the sum of the n downstream processing capacities used to produce product j in refinery i at time t, L_{sjt} is the labor used in state s to produce product j at time t, and $\boldsymbol{Z}_{ijt}$ is a vector of additional inputs and other control variables including additional fuel additives, crude oil quality, and market share.[27] Note that because I am estimating a multi-product production function, unobserved productivity, Ω_{ijt}, is indexed by refinery, product, and

[27] Refineries do not report labor or energy usage. I therefore use state level labor input estimates from the Bureau of Labor Statistics. Although this may lead to omitted variable bias, the proportion of energy and labor used in the production process is relatively small. According to the Energy Information Administration, energy and labor account for only about 2.7% and 4.2% of the production costs respectively. As a robustness check I tried a number of different specifications for labor including weighting the labor variable by refinery within-state capacity share. None of the specifications had significant effects on the crude oil output elasticity estimates, the coefficient of interest.

time.[28] This means in the estimation procedure I must also address any unobserved aspect of input allocation choices. Finally, $\boldsymbol{\beta}$ is a vector of parameters to be estimated.

Taking the log of equation (4) yields

$$q_{ijt} = f\left(c_{ijt}, k_{ijt}, dk_{ijt}, l_{sjt}, \boldsymbol{z}_{ijt}; \boldsymbol{\beta}\right) + \omega_{ijt} + \epsilon_{ijt}, \tag{5}$$

where lowercase letters represent logged variables. Let $\omega_{ijt} = \ln(\Omega_{ijt})$ and ϵ_{ijt} represent log additive measurement error or unexpected shocks to output.

Estimating a multi-product production function poses two primary identification concerns. First, unobserved productivity, ω_{ijt}, has the potential to cause simultaneity and selection biases. For instance, some firms may be relatively more efficient than others, may have regional market power, or may have access to better production technologies. Each of these unobserved factors could lead to correlation between input choices and unobserved productivity.

The second endogeneity concern can be made explicit by rewriting the production function (equation (5)) as follows:

$$q_{ijt} = f(\tilde{\boldsymbol{x}}_{ijt}, \boldsymbol{z}_{ijt}; \boldsymbol{\beta}) + \omega_{ijt} + \epsilon_{ijt}, \tag{6}$$

where $\tilde{\boldsymbol{x}}_{ijt} = \log\left(\rho_{ijt}\boldsymbol{X}_{it}\right)$, and $\boldsymbol{X}_{it}$ is the vector of observed refinery level inputs while ρ_{ijt} represents the share of inputs devoted to producing product j, or the input allocation. The concern is that refiners may choose input allocation endogenously, often referred to as effort. Intuitively, a firm will exert more effort to produce a given product if it can charge higher markups for that product. The input allocation is therefore potentially correlated with unobserved demand conditions, seasonality, and market power. Any unobserved component of input allocation is also captured in ω_{ijt} so the two identification concerns can be addressed simultaneously.

Estimation Procedure and Identification

Employing a translog functional form to estimate (5) results in the following empirical specification

$$q_{ijt} = (1 + \tilde{dk}_{ijt})\beta_{dk} + \boldsymbol{z}_{ijt}\beta_z + g\left(\tilde{c}_{ijt}, \tilde{k}_{ijt}, \tilde{l}_{sjt}\right) + \omega_{ijt} + \epsilon_{ijt}, \tag{7}$$

where $g(\cdot)$ contains the translog terms (and the corresponding coefficients) including all primary inputs $\left(\tilde{c}_{ijt}, \tilde{k}_{ijt}, \tilde{l}_{sjt}\right)$, primary inputs squared, and interaction terms between all primary inputs. I assume that downstream capacity is multiplicative because it can increase the proportion of a given end product extracted from a given unit of input.

[28] As is common in the literature, I assume that the Hicks-neutral productivity, Ω_{ijt}, is log-additive and firm-product-specific. Hicks neutral productivity simply means that any technological change does not affect the marginal rate of substitution between any two inputs. This assumption allows the productivity term to be multiplicative and therefore invertible, ultimately allowing estimation of the production function using a proxy approach, e.g. Olley and Pakes (1996) and Levinsohn and Petrin (2003).

Other control variables are assumed to enter linearly and the intercept is subsumed in ω_{ijt}. A benefit of the translog specification is that it allows the output elasticities to vary across firms and products. As De Loecker and Warzynski (2012) point out, restricting output elasticities to be constant across firms and products when computing markups would attribute variation in technology to variation in markups.

I employ the proxy method developed by Olley and Pakes (1996) (OP) and built upon by Levinsohn and Petrin (2003) (LP) and Ackerberg, Caves, and Frazer (2015) (ACF) to address potential bias associated with unobserved productivity and effort and to obtain consistent estimates of θ^c_{ijt}.[29] The proxy method begins by defining the intermediate input demand function. Assume crude oil input demand for product j is a function of productivity, current period capital and labor, and a vector $\boldsymbol{m}_{ijt} = \{ms_{ijt}, RFS_t, \tilde{dk}_{ijt}\}$, which includes downstream capital, market share for each product, and the renewable fuel credit price RFS_t,

$$\tilde{c}_{ijt} = c_t(\omega_{ijt}, \tilde{k}_{ijt}, \tilde{l}_{sjt}, \boldsymbol{m}_{ijt}). \tag{8}$$

Following ACF, labor is included in (8), implying that labor choices affect crude input demand. The additional control variables are included to control for factors affecting input demand choices across firms. For instance, in Section 9, I show that the RFS credit price affects markups and therefore input choices. Anticipating this result, I include the RFS credit price in the input demand function to control for its affect on the output elasticities. The intermediate input demand function is then inverted to form an expression for unobserved productivity as a function of crude oil inputs, capital, labor, and the variables collected in $\boldsymbol{m}_{ijt}$,[30]

$$\omega_{ijt} = h_t\left(\tilde{c}_{ijt}, \tilde{k}_{ijt}, \tilde{l}_{sjt}, \boldsymbol{m}_{ijt}\right). \tag{9}$$

Plugging (9) into (7) results in the first stage equation

$$q_{ijt} = (1 + \tilde{dk}_{ijt})\beta_{dk} + \boldsymbol{z}_{ijt}\beta_z + \phi_t(\tilde{c}_{ijt}, \tilde{k}_{ijt}, \tilde{l}_{sjt}, \boldsymbol{m}_{ijt}) + \epsilon_{ijt}, \tag{10}$$

where

$$\phi_t(\cdot) = g\left(\tilde{k}_{ijt}, \tilde{c}_{ijt}, \tilde{l}_{sjt}\right) + h_t\left(\tilde{c}_{ijt}, \tilde{k}_{ijt}, \tilde{l}_{sjt}, \boldsymbol{m}_{ijt}\right), \tag{11}$$

and $\boldsymbol{z}_{ijt}$ contains crude oil quality, market share, and additional fuel specific additives such as oxygenates and renewable fuels.

Estimation proceeds in two stages. In the first stage I estimate equation (10) by

[29] In contrast, to obtain consistent estimates of θ^c_{ijt} using a reduced form approach would require a minimum of one variable that shifts crude oil input choices and is uncorrelated with productivity. Input prices are an obvious choice but are associated with a host of well known issues (Ackerberg et al. 2007). Likewise, to obtain consistent estimates from a cost function, the dual to the production function, I would need a refinery specific instrument to shift input choices such as a refinery specific exogenous demand side variable. However, such data is not readily available in most contexts. A further disadvantage of the reduced form approach in the present context is that I observe crude oil prices at the firm level, not the refinery level. This means that the reduced form production function estimation must also be at the firm level.

[30] Inversion assumes $\frac{\partial \tilde{c}_{ijt}}{\partial \omega_{ijt}} > 0$ conditional on $\boldsymbol{m}_{ijt}$. The monotonicity of intermediate inputs with respect to productivity changes has been shown to hold under imperfect competition (Melitz and Levinsohn 2006 see De Loecker and Warzynski for the reference). The monotonicity assumption has been proven for the Cournot case in which higher productivity firms (low marginal cost) must use higher quantities of intermediate inputs at any level of residual demand (De Loecker 2011).

replacing $\phi_t(\cdot)$ with a nonparametric polynomial expansion of capital, labor, and crude oil inputs while the remaining control variables enter linearly.[31] Following ACF, I diverge from the traditional LP approach by including the primary dynamic and variable inputs from the intermediate demand function in $\phi_t(\cdot)$. Doing so, I obtain coefficient estimates for market share, downstream capacity, crude oil quality, and additional fuel additives in the first stage nonparametric regression. I also obtain an estimate of $\hat{\phi}_t(\cdot)$, net of the unexpected error term, ϵ_{ijt}.

The second stage of estimation identifies the translog coefficients in $\boldsymbol{\beta}$ by relying on assumptions about the evolution of productivity, and the timing of input use relative to productivity shocks. Assuming productivity evolves via a first order Markov process allows current period productivity to be expressed as a function of lagged productivity, the RFS credit price, and an unexpected deviation, ξ_{ijt}, as follows[32]

$$\omega_{ijt} = g_t\left(\omega_{ijt-1}, \ln(RFS_{t-1})\right) + \xi_{ijt}. \tag{12}$$

The unexpected deviation term, ξ_{ijt}, captures random shocks to productivity and may also contain information about shocks to a refinery's effort or input allocation choice. The lag of the RFS credit price is included in the law of motion because changes in the RFS credit price are shown to affect firm level markups and therefore competition.

An estimate of productivity, for any vector of translog coefficients $\boldsymbol{\beta}_{TL}$, can then be expressed as

$$\begin{aligned}
\omega_{ijt}(\beta_j) \quad = \quad & \hat{\phi}_{ijt}(\tilde{c}_{ijt}, \tilde{k}_{ijt}, \tilde{l}_{sjt}, \boldsymbol{z}_{ijt}) && (13)\\
& -\beta_k\tilde{k}_{ijt} - \beta_l\tilde{l}_{sjt} - \beta_c\tilde{c}_{ijt} - \beta_{kk}\tilde{k}^2_{ijt} - \beta_{cc}\tilde{c}^2_{ijt} - \beta_{ll}\tilde{l}^2_{sjt} && (14)\\
& -\beta_{ck}\tilde{c}_{ijt}\tilde{k}_{ijt} - \beta_{cl}\tilde{c}_{ijt}\tilde{l}_{sjt} - \beta_{kl}\tilde{k}_{ijt}\tilde{l}_{sjt} - \beta_{ckl}\tilde{c}_{ijt}\tilde{k}_{ijt}\tilde{l}_{sjt}. && (15)
\end{aligned}$$

Finally, by nonparametrically estimating equation 12, I can recover an estimate of the productivity shock, ξ_{ijt}.[33]

Identification depends crucially on explicit assumptions about the relationship between a refinery's current period information set and the timing of the observed inputs choices. I assume that capital and labor are dynamic inputs, meaning current choices of capital and labor affect future profits and importantly, current period capital and labor were chosen in $t-1$ or earlier. In addition, adjustments to capital and labor are assumed to be costly and take time to complete, e.g., investments in capital in period t will not be reflected in the production process until at least $t+1$. Thus, current period random productivity shocks are uncorrelated with current primary capital, downstream capital, and labor choices, $E[\xi_{ijt}\tilde{k}_{ijt}] = E[\xi_{ijt}\tilde{dk}_{ijt}] = E[\xi_{ijt}\tilde{l}_{sjt}] = 0$. These are the primary moment conditions used

[31] Specifically, $\phi_t(\cdot)$ contains the squared and cubed terms of each of the variable and dynamic input variables and all of the interactions between the squared and level terms.

[32] An additional assumption noted by ACF is that ω_{ijt} must be the only unobservable entering the refinery's intermediate input demand function. This assumption implies that any unobserved relationship between a refinery's choice of inputs or a refinery's choice of input allocation must be captured in ω_{ijt}.

[33] In practice, (12) is estimated using $\omega_{ijt} = \alpha_0 + \alpha_1\omega_{ijt-1} + \alpha_2\omega^2_{ijt-1} + \alpha_3\omega^3_{ijt-1} + RFS_{t-1} + \nu_{ijt}$.

to identify the capital and labor coefficients. In contrast, I assume crude oil is a freely adjustable input and as such, the refinery chooses current period inputs in period t after observing ω_{ijt}. This implies that the choice of crude oil inputs today may be correlated with current period innovation, $E[\xi_{ijt}\tilde{c}_{ijt}] \neq 0$. However, since crude inputs are chosen in each period, lagged inputs are uncorrelated with current period productivity shocks implying $E[\xi_{ijt}\tilde{c}_{ijt-1}] = 0$, which is the primary moment condition that will be used to identify the coefficient on crude inputs.[34]

Input allocation choices are also potentially correlated with unobserved market power and other potentially unobserved variables. However, refineries likely make current period crude oil input choices, as well as input allocation choices, based on their market share history. As such, lagged market share is used as an additional instrument. I control for endogeneity in input allocation choices by including lagged market share and lagged market share interacted with crude oil, capital, and labor choices in the instrument vector. Current period downstream capacity is also included in the instrument vector for additional product level identification.

Using these identification assumptions, I form the following moments to estimate the production function

$$E\left(\xi_{ijt}(\boldsymbol{\beta}_{TL})\boldsymbol{Y}^h_{ijt}\right) = 0, \tag{16}$$

where $\boldsymbol{Y}^h_{ijt}$ is a vector of instruments indexed by h, containing current period capital and labor, lagged crude oil inputs, current downstream capacity, and all associated translog terms, lagged market share, and the interaction of lagged market share with current capital, current labor, lagged crude oil inputs, and current period downstream capacity.[35] I used standard GMM techniques to recover the translog coefficients, $\boldsymbol{\beta}_{TL} = [\beta_l, \beta_k, \beta_c, \beta_{ll}, \beta_{kk}, \beta_{cc}, \beta_{cl}, \beta_{kl}, \beta_{kc}, \beta_{ckl}]$. As is common in the literature, standard errors are bootstrapped and clustered at the refinery level. In practice this means drawing a random sample of refineries with replacement.[36]

Obtaining Markups from Output Elasticities and Data

Using the translog functional form means the estimate of the output elasticity with respect to crude oil inputs is obtained from the estimated production function coefficients via the following expression[37]

[34]In light of evidence that refineries face relatively substantial adjustment costs (Borenstein and Shepard 2002), the assumption that crude oil is a freely adjustable input may be controversial. As noted in section (4), refineries make production runs in which machinery is tuned specifically for a given grade of crude and for a particular output mix. Since production runs typically last 2-4 weeks, the level of observation in my data is sufficient to assume firms can freely adjust crude inputs.

[35]I do not use crude oil input prices as an additional source of variation because crude oil input prices are observed at the firm level rather than the refinery level. However, when crude oil input prices are used as instruments, the coefficients do not change much.

[36]For more information see Appendix A.

[37]A note on the interpretation of the parameter of interest, θ^c_{ijt}, is warranted to clarify a potential point of confusion. One might be concerned that estimating the production function from (6) would provide an estimate of the average percentage of a given end product produced from a barrel of oil, call this value γ. For example, the California Energy Commissions reports that approximately 51.4% of a barrel of oil is converted into gasoline on average. If one were to estimate (6)

$$\hat{\theta}^c_{ijt} = \hat{\beta}_c + 2\hat{\beta}_{cc}\tilde{c}_{ijt} + \hat{\beta}_{cl}\tilde{l}_{sjt} + \hat{\beta}_{ck}\tilde{k}_{ijt} + \hat{\beta}_{ckl}\tilde{l}_{sjt}\tilde{k}_{ijt}, \tag{17}$$

where hats represent the second stage translog coefficient estimates.

I aggregate refinery level crude oil inputs and end product outputs to the firm-PADD level, $\tilde{C}_{frjt} = \sum_{i \in f^r} \tilde{C}_{ijt}$ and $Q_{frjt} = \sum_{i \in f^r} Q_{ijt}$ respectively, to estimate the markups at the firm-product-PADD level, where PADD is indicated by the letter r. I observe firm-PADD specific crude oil input prices, p^c_{frt}, and firm-product-state specific wholesale output prices, P_{fsjt}. I average the wholesale output prices to the firm-PADD level, P_{frjt}. Finally, I average the refinery level output elasticities to the firm-PADD level, $\hat{\theta}^c_{frjt}$. The sample estimate of markups follows from equation (2)

$$\hat{\mu}_{frjt} = \hat{\theta}^c_{frjt}\left(\frac{P_{frjt}Q_{frjt}}{p^c_{frt}\tilde{C}_{frjt}}\right). \tag{18}$$

Variation in markups comes from variation in output prices across products, firms, and regions, input prices across firms and regions, output elasticities across products, firms, regions, and the input to output ratio across firms and regions. An additional benefit of estimating markups in this fashion is that I can also recover an estimate of marginal costs for each firm product combination by rearranging the markup term,

$$\hat{mc}_{frjt} = \frac{P_{frjt}}{\hat{\mu}_{frjt}}.$$

8 Production Function, Markup, and Marginal Cost Estimation Results

In the following section I provide the results of estimating a refinery-product level production function using the ACF routine, and the resulting markup and marginal cost estimates.

Production Function Results

Table 4 presents the results of estimating the production function, equation (7). Rather than reporting each of the translog coefficients, I report the output elasticities and their bootstrapped standard errors. The coefficients are generally within the range of estimates in the literature.

Not surprisingly, the average output elasticities with respect to crude oil inputs for gasoline and diesel are not significantly different from one another. This is likely due to the similarity in the production processes across all fuels, and the close relationship

omitting ρ_{ijt}, then this would indeed provide an estimate of γ. In other words, γ estimates the average percentage of crude converted into gasoline conditional on the quantity of crude input and other control variables, $\gamma = E(\frac{Q_{ijt}}{Cit}|C_{it})$ where C_{it} is refinery $i's$ crude inputs in period t. In contrast, I estimate $\theta^c_{ijt} = \frac{\partial \ln Q_{ijt}}{\partial \ln\left(C_{it} * (Q_{ijt}/\sum_j Q_{ijt})\right)}$.

Table 4: ACF Production Function Results

	Gas	Jet	Diesel	Low End
Crude Input (θ^c_{ijt})	0.859***	0.791***	0.872***	0.973***
	(0.019)	(0.011)	(0.044)	(0.025)
Capacity (θ^k_{ijt})	0.104***	0.178***	0.055***	0.015
	(0.020)	(0.013)	(0.012)	(0.026)
Labor (θ^l_{ijt})	0.020***	0.006	0.079	0.002
	(0.005)	(0.014)	(0.05)	(0.003)
Market Share	0.030***	0.022***	0.020***	0.014***
	(0.001)	(0.001)	(0.000)	(0.000)
Sulfur Content	0.161***	0.023	0.048	0.144***
	(0.029)	(0.039)	(0.036)	(0.023)
API Gravity	-0.005***	-0.007***	0.008***	0.010***
	(0.001)	(0.002)	(0.002)	(0.001)
Downstream Capacity	0.002**	0.007***	0.007***	0.002***
	(0.001)	(0.001)	(0.001)	(0.000)
Oxygenates and UFO	-0.001			
	(0.000)			
Kerosene UFO Inputs		-0.002***		
		(0.000)		
Renewable Inputs	0.012***		0.020***	
	(0.000)		(0.001)	
N	26355	12087	15401	38342

Notes: Products include gasoline, diesel, jet fuel, and low-end products. The coefficient estimates reported for crude input, capcity, and labor are the output elasticity estimates generated using the translog output elasticity expression. Standard errors are bootstrapped using the procedure described in Appendix A. All other coefficient estimates are from the first stage of the ACF estimation routine.
***Significance at the 1 percent level.
**Significance at the 5 percent level.
*Significance at the 10 percent level.

between crude oil inputs and end product outputs. The output elasticities for jet fuel and low end products tend to be the lowest and highest respectively while the output elasticities for gasoline and diesel fall somewhere in the middle. The distribution of estimated crude oil output elasticities with respect to gasoline across firms is fairly tight, ranging from approximately 0.83 to 0.89.

Additional fuel additives and downstream capacity appear to be extremely important. When these additional control variables are removed, the crude oil output elasticity becomes significantly higher. This highlights the importance of controlling for intermediate inputs and capacity.

I perform a number of robustness checks to provide confidence in the output elasticity results and to test my input allocation assumption. First, I estimate a Cobb-Douglas production function. The Cobb-Douglas estimates for gasoline, diesel, and jet fuel are not statistically different from one another, which indicates the importance of including variation in input use intensity when estimating the output elasticities and allowing the output elasticities to vary across firms and products. Second, I estimate a reduced form firm-product level production function, instrumenting crude oil input choices with firm level input prices. The crude oil output elasticity estimate is generally higher using two stage least squares than using the ACF approach, with the exception of the low end products elasticity. Controlling for unobserved productivity using the ACF reduces the elasticity estimate because productivity is positively correlated with input choice, except in the case of low end outputs. Differences between the reduced form and ACF estimates can also be attributed to the level of estimation (firm or refinery level) and the method used to control for endogeneity. Where the reduced form approach assumes productivity is constant over time, the ACF approach explicitly allows productivity to evolve. Indeed, previous research shows an increasing trend in estimated productivity over time and that refinery productivity increased in California in response to regulation (Berman and Bui 2001).

Ideally I could also test my input allocation assumption by estimating the production function using revenue shares in place of physical output shares. However, I observe output prices for only a subset of the products each refinery produces so I cannot calculate a refinery's total revenue.[38]

Markup and Marginal Cost Results

The distributions of estimated markups and marginal costs are presented in Figures 4 and 5 respectively. Average markups are 1.23 with a standard error of 0.016.[39] This implies that petroleum product prices are 23% greater than marginal costs on average.[40] Comparing the two distributions highlights an important feature of the estimation strategy: markups and marginal costs are not necessarily inversely related to one another and may move in the same or different directions.

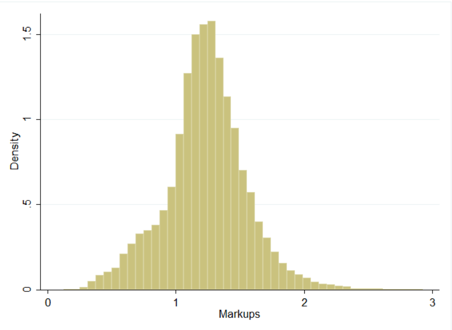

Figure 4: Distribution of Estimated Markups

Notes: Inner 98th Percentile.

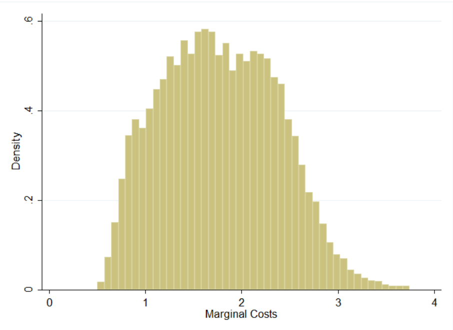

Figure 5: Distribution of Estimated Marginal Costs

Notes: Inner 98th Percentile.

Table 5 provides markup summary statistics by PADD. There are a large number of pipelines linking the Gulf Coast (PADD 3) to the Eastern Seaboard (PADD 1) and

[38]Robustness results are available upon request.
[39]Standard errors are bootstrapped. See Appendix A for details.
[40]While 95% of the markups are below 1.75, I estimate a maximum markup of 543. Obviously firms do not have a markup 543 times greater than marginal cost. Outliers such as this occur because of random variation in the output and crude oil input variables, which could be due to either data recording errors or spills for example. I therefore truncate the markup and marginal cost estimates removing estimates below the 1st and above the 99th percentiles.

the Midwest (PADD 2), and to some extent the Rocky Mountain Region (PADD 4). However the West Coast (PADD 5) is quite isolated due to the natural barriers of the Rocky and the Sierra Mountain Ranges. Moreover, California has additional content regulations above and beyond the national standards. It is therefore not surprising that average markups are relatively high along the West Coast. In contrast, average markups are lowest in the Gulf Coast implying competition is highest in that region.

Table 5: Markup Summary Statistics By Region

PADD	**Mean**	**Std. Dev.**	**Min.**	**Max.**	**N**
East Coast (1)	1.272	0.020	0.325	2.801	3,671
Midwest (2)	1.243	0.016	0.226	2.732	8,854
Gulf Coast (3)	1.187	0.017	0.336	2.927	8,982
Rocky Mountain (4)	1.228	0.021	0.28	2.861	4,693
West Coast (5)	1.272	0.017	0.124	2.91	6,305
Total	1.234	0.016	0.124	2.927	32,505

Notes: Summary statistics are for the inner 98th percentile of estimated markups. Markups can be interpreted as the percentage markup over marginal costs, i.e., 1.273 represents a 27% markup over marginal costs.

Table 6: Markup Summary Statistics By Fuel Type

Fuel	**Mean**	**Std. Dev.**	**Min.**	**Max.**	**N**
Conventional Gas	1.29	0.023	0.392	2.776	6,371
Reformulated Gas	1.316	0.023	0.418	2.817	3,863
Regular Diesel	1.478	0.059	0.661	2.927	3,181
ULSD	1.413	0.060	0.425	2.861	5,269
Jet Fuel	1.197	0.022	0.336	2.647	4,016
Low End Products	1.006	0.013	0.124	2.601	9,805
Total	1.234	0.016	0.124	2.927	32,505

Notes: Summary statistics are for the inner 98th percentile of estimated markups. Markups can be interpreted as the percentage markup over marginal costs, i.e., 1.273 represents a 27% markup over marginal costs.

Table 7: Marginal Cost Summary Statistics By Region ($/gallon)

PADD	**Mean**	**Std. Dev.**	**Min.**	**Max.**	**N**
East Coast (1)	1.761	0.024	0.42	4.87	3,671
Midwest (2)	1.749	0.022	0.497	3.723	8,854
Gulf Coast (3)	1.836	0.026	0.296	4.957	8,982
Rocky Mountain (4)	1.67	0.027	0.545	3.612	4,693
West Coast (5)	1.828	0.021	0.475	8.122	6,305
Total	1.778	0.021	0.296	8.122	32,505

Notes: Summary statistics are for the inner 98th percentile of estimated marginal costs.

Average markups by region are conditional on the products sold in those regions. Looking at markups by product (Table 6), I find that firms charge higher markups for conventional diesel than ultra-low-sulfur diesel. Consistent with previous research, I find that firms charge higher markups for reformulated gasoline than conventional gasoline

(Sweeney 2015). Finally, I find that some firms actually charge below marginal cost for their residual fuels.

Marginal cost summary statistics by region and fuel type are reported in Tables 7 and 8 respectively. Average marginal costs across all fuel types are highest in the Gulf Coast and West Coast regions, consistent with crude oil input prices in these regions (Table 9). Variation in crude oil input prices may be attributed to variation in crude oil quality (the West Coast processes higher quality crude) and access to low price domestic crude oil post 2011. Indeed the Rocky Mountain region has the lowest marginal costs and crude oil input prices on average. Marginal costs by fuel type are reported in Table 8. On average, gasoline marginal costs are higher than conventional diesel while jet fuel and ultra-low-sulfur diesel have the highest marginal costs. It is important to note that some of the variation in marginal costs across products may be attributed to where these products are made rather than their respective production processes. For instance, the majority of the jet fuel produced in the U.S. may come from regions with high crude oil input prices such as the Gulf Coast.

Table 8: Marginal Cost Summary Statistics By Fuel Type ($/gallon)

Fuel	**Mean**	**Std. Dev.**	**Min.**	**Max.**	**N**
Conventional Gas	1.828	0.043	0.508	7.434	6,371
Reformulated Gas	1.85	0.042	0.509	7.456	3,863
Regular Diesel	1.531	0.069	0.441	4.957	3,181
ULSD	1.863	0.095	0.517	6.764	5,269
Jet Fuel	2.048	0.040	0.601	8.122	4,016
Low End Products	1.642	0.018	0.296	6.624	9,805
Total	1.778	0.021	0.296	8.122	32,505

Notes: Summary statistics are for the inner 98th percentile of estimated marginal costs.

Table 9: Crude Oil Input Price Summary Statistics By Region ($/gallon)

PADD	**Mean**	**Std. Dev.**	**Min.**	**Max.**	**N**
East Coast (1)	1.774	0.630	0.681	3.254	3,671
Midwest (2)	1.761	0.559	0.593	3.254	8,854
Gulf Coast (3)	1.876	0.609	0.659	3.278	8,982
Rocky Mountain (4)	1.682	0.511	0.657	3.162	4,693
West Coast (5)	1.855	0.596	0.660	3.242	6,305
Total	1.801	0.586	0.593	3.278	32,505

Notes: Summary statistics for firm reported crude input prices by region.

To provide additional confidence that the markups behave in an intuitive fashion, I relate estimated markups to capacity and market share. As expected, markups are increasing with a firm's PADD-level market share (Figure 7). Additionally, Figure 6 shows a positive relationship between estimated markups and observed capacity. While Texas and Louisiana boast the highest concentration of refineries in the U.S., the refineries in these states are also some of the largest. As previously mentioned, there are proportional production gains associated with refining crude oil, which implies that larger refineries are more efficient and can charge higher markups on average.[41]

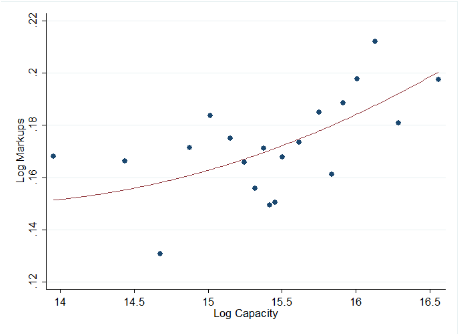

Figure 6: Markups and Atmospheric Distillation Capacity

Notes: Absorbing firm, fuel, PADD, and time fixed effects.

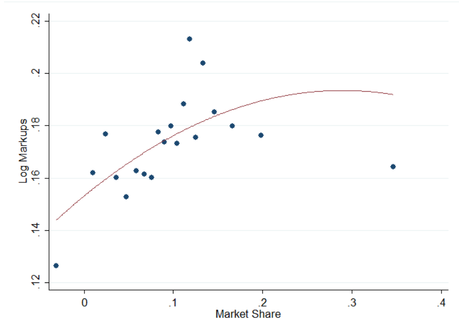

Figure 7: Markups and Market Share

Notes: Absorbing firm, PADD, and time fixed effects.

9 The Renewable Fuel Standard and Refinery Prices, Markups, Marginal Costs, and Production Decisions

With markups and marginal costs in hand, I can now turn to evaluating the effect of changes in the RFS credit price on refinery behavior - the main focus of this paper. I evaluate the effect of the RFS on four separate variables: wholesale output prices, markups, marginal costs, and product shares.

Output Prices and Renewable Fuel Credit Prices

I begin by evaluating the relationship between wholesale petroleum product prices and the RFS credit price. Similar to Knittel, Meiselman, and Stock (2015), I estimate the long run pass-through relationship by estimating level regressions.[42] Specifically, I estimate

$$P_{frjt} = \beta_0 + \beta_1 RFS_t + \beta_2 p^c_{frt} + \sigma_{fy} + G_r + S_s + J_j + \epsilon_{fjst}, \tag{19}$$

where P_{frjt} is the output price charged by firm f in region r for product j at time t (less transportation costs), RFS_t is the RFS credit price at time t, p^c_{frt} is the average price per gallon of crude oil for firm f in region r, and σ_{fy}, G_r, S_s, J_j are firm-year, region, season, and product fixed effects respectively. I control for firm-region specific crude oil prices as the majority of the movement in wholesale prices can be attributed to movements in the price of crude oil. Standard errors are clustered at the firm-month level as the credit price varies at the monthly level while the errors may be correlated within firms.

The results of estimating (19) are presented in Table 10. Column 1 shows the average relationship between regulated fuel prices (gasoline and diesel) and the RFS credit price for 2011-2012. Prior to 2013, the RFS credit price was negligible and fairly stable.

[41]Refineries report production gains and losses on survey form EIA-810.
[42]First difference regressions provide fairly similar results.

Hence, in 2011 and 2012 there was not enough variation in the RFS credit price to detect a significant effect on output prices. On the other hand, column 2 shows the relationship between regulated fuel prices and the RFS credit price for 2013 and 2014 only; a period in which policy uncertainty caused a series of significant price shocks and volatility (Refer to Section 4 for details). During this time period I estimate a coefficient of 0.93 meaning that a one dollar increase in the RFS credit price resulted in a ninety three cent increase in gasoline and diesel prices on average. Columns 3-6 break this relationship down by fuel type. Column 3 shows that RFS credit prices were excessively passed-through in the gasoline market. In contrast, column 5 shows a less than complete pass-through rate for ultra-low-sulfur diesel while the credit price appears to have had no statistically significant effect on regular diesel prices (column 4). The difference in statistical significance between the regular and ultra-low-sulfur diesel coefficients could reflect the fact that the two fuels are sold in different markets. Ultra-low-sulfur diesel is the primary fuel used on the road in trucks and cars while regular diesel is sold for commercial and industrial uses. Another important piece of information in Table 10 is the low pass-through rate of crude oil prices. Prior to 2011, I estimate an average pass-through rate of 0.96. However, domestic crude oil production dramatically increased in 2011, which, combined with an oil export ban caused a large spread in the domestic and international crude oil prices. After 2011, refineries paid very different input prices depending on their access to cheap domestic oil. Hence, output prices for refined products were less sensitive to variation in crude oil prices post 2011.

Table 10: Output Prices and the RFS Credit Price

	G and D 2011-12 (1)	G and D ≥2013 (2)	Gas ≥2013 (3)	Diesel ≥2013 (4)	ULSD ≥2013 (5)	Jet ≥2013 (6)
RFS Credit Price	-0.239	0.931***	1.248***	0.408	0.612***	-1.050***
	(0.915)	(0.178)	(0.231)	(0.359)	(0.134)	(0.206)
Crude Price	0.324***	0.199***	0.264***	0.107*	0.170***	0.171***
	(0.021)	(0.024)	(0.030)	(0.054)	(0.019)	(0.023)
Firm-Year FE	Y	Y	Y	Y	Y	Y
Product FE	Y	Y	Y	NA	NA	NA
Seasonal FE	Y	Y	Y	Y	Y	Y
Region FE	Y	Y	Y	Y	Y	Y
R-squared	0.620	0.555	0.573	0.519	0.368	0.513
N	3651	3085	1540	487	1058	796

Notes: Column 1 and 2 pool gasoline and diesel prices. The dependent variable in columns 3-6 are gasoline, diesel, ultra-low-sulfur diesel, and jet fuel respectively. Column 1 uses data from 2011-2012 only while columns 2-6 use only 2013-2014 data. Standard errors are clustered at the firm-month.
***Significance at the 1 percent level.
**Significance at the 5 percent level.
*Significance at the 10 percent level.

Column 6 of Table 10 shows that non-regulated fuel prices were affected as well. Jet fuel prices decreased in response to increases in the RFS credit price. The effect on jet fuel prices may be attributed to increased production of jet fuel causing jet fuel prices to fall.[43]

The relationship between the RFS credit price and jet fuel prices is important. Knittel, Meiselman, and Stock (2015) estimate the pass-through rate of the RFS credit price to 2013-2015 gasoline and diesel spot prices using jet fuel prices, and in some cases crude oil spot prices, as a control for aggregate movements in petroleum product prices. The evidence presented here shows that price trends in non-regulated fuels in a multi-product setting do not necessarily provide adequate controls for regulated fuel price trends. I estimate pass-through rates similar to those estimated in Knittel, Meiselman, and Stock (2015) without using jet fuel prices as a control. Instead, I control for average trends in output prices by including the firm specific crude oil input prices and seasonal dummies as independent variables. Knittel, Meiselman, and Stock (2015) estimate a long run pass-through rate of 1.01 on average across diesel and gasoline spot prices with considerable variation at the daily and weekly level. However, they consistently estimate greater than 100% long run pass-through in the diesel market. Looking at 2013 only, they estimate much higher pass-through rates, in some cases as high as 4.299. Our results combined suggest there is significant variation in the pass-through rate of the RFS credit price to product prices across time periods, fuels, and price trends.

Markups, Marginal Costs, and Renewable Fuel Credit Prices

Next, I relate the RFS credit prices to markups and marginal costs using a simple regression framework. Doing so allows me to decompose the pass-through results presented in the previous section. The regression equation is

$$Y_{frjt} = \delta_0 + \delta_1 \ln RFS_t + \delta_2 \boldsymbol{X}_{frt} + \sigma_{fy} + G_r + S_s + J_j + \varepsilon_{fjst}, \tag{20}$$

where Y_{frjt} is one of two dependent variables: the log of markups for firm f in region r and product j at time t ($\ln \mu_{frjt}$), or the log of marginal costs for firm f in region r producing product j at time t ($\ln mc_{frjt}$). The variable of interest, $\ln RFS_t$, is the log of the RFS credit price for period t, and $\boldsymbol{X}_{frt}$ is a vector of control variables including the number of firms in the market, a firm's market share for product j, firm level productivity, and the log of firm specific crude oil prices.[44] The remaining variables are firm-year, state, seasonal, and product fixed effects. The parameter δ_1 captures the effect of a one percent change in the RFS credit price on a firm's marginal costs or markups. Bootstrapped standard errors are generated by iterating over the entire production function estimation routine (See Appendix A for details).

Marginal cost results for gasoline (conventional and reformulated), diesel, and jet fuel are presented in columns 1, 4, and 6 of Table 11 respectively. The results suggest that

[43]Estimating equation (19) with varying combinations of fixed effects has little effect on the coefficient estimates with the exception of the seasonal fixed effect. Omitting seasonal fixed effects increases the magnitude of the coefficients but does not change the sign. Omitting crude oil prices has a similar effect.

[44]Markets are defined at the PADD-level.

between 2013 and 2014, a 10% (0.7 ¢/gallon) increase in the RFS credit price increased gasoline and diesel marginal costs by approximately 0.08% (0.2 ¢/gallon) and 0.18% (0.4 ¢/gallon) respectively, conditional on a firm's productivity, market share, crude oil input prices, and the market concentration. Interestingly, I also find that jet fuel marginal costs, a non-regulated product, increased in response to the RFS credit price. This result can be attributed to the joint nature of the petroleum refining process, and will be explored further below.

Table 11: Marginal Costs, Markups, and the RFS Credit Price

	Log MC Gas (1)	Log μ Conv G (2)	Log μ Reform G (3)	Log MC Diesel (4)	Log μ Diesel (5)	Log MC Jet (6)	Log μ Jet (7)
Log RFS	0.008***	0.036***	0.025***	0.018***	0.006	0.016**	-0.028***
Credit Price	(0.003)	(0.005)	(0.004)	(0.005)	(0.005)	(0.007)	(0.006)
Market Share	-0.094**	0.201***	-0.011	-0.001	-0.062**	-0.261***	0.287***
	(0.047)	(0.056)	(0.087)	(0.025)	(0.025)	(0.080)	(0.087)
Log Crude Price	0.907***	-0.705***	-0.705***	0.943***	-0.845***	0.842***	-0.741***
	(0.034)	(0.038)	(0.058)	(0.034)	(0.035)	(0.043)	(0.044)
Firm-Year FE	Y	Y	Y	Y	Y	Y	Y
Product FE	Y	N/A	N/A	Y	Y	N/A	N/A
Region FE	Y	Y	Y	Y	Y	Y	Y
Seasonal FE	Y	Y	Y	Y	Y	Y	Y
Controls	Y	Y	Y	Y	Y	Y	Y
R-squared	0.727	0.661	0.671	0.761	0.754	0.66	0.644
N	1558	955	603	1262	1262	743	743

Notes: These results are for 2013-2014. The dependent variables are either log marginal costs or log markups for all gasoline, conventional gasoline, reformulated gasoline, all diesel, or jet fuel. Control variables include market share, productivity, crude oil prices, and the number of firms in the market. Standard errors are bootstrapped and clustered at the refinery level. See Appendix A for details.
***Significance at the 1 percent level.
**Significance at the 5 percent level.
*Significance at the 10 percent level.

The effect of the RFS on markups is also reported in Table 11. To capture potential competitive effects of changes in the RFS credit price, I need to control for simultaneous shocks to marginal costs (De Loecker et al. 2016; De Loecker and Warzynski 2012). It is well known that the primary and most volatile cost of refining is the cost of crude oil. Therefore, to control for short term changes in marginal costs, I include firm specific crude oil input prices. To control for longer term variation in marginal costs, I include firm specific productivity, which I estimate from the production function routine.[45] As expected, including these additional control variables reduces the estimate of δ_1.

Columns 2, 3, 5, and 7 of Table 11 present the results of regressing the log of conven-

[45] I do not control for marginal costs directly because the RFS credit price is highly correlated with the marginal cost estimates. However, including marginal costs rather than crude input prices does not change the sign or the significance of the coefficients. As noted by De Loecker and Warzynski (2012), controlling for productivity implies controlling for differences in marginal cost across firms. This eliminates the productivity component from the markup estimates which allows the researcher to isolate the role of other factors that impact prices not included in firm level productivity, i.e., differences in elasticities of demand across markets and products. The coefficient on productivity will pick up potential variation across firms including market power and demand conditions. See De Loecker and Warzynski (2012) for a detailed description of this relationship.

tional gasoline, reformulated gasoline, diesel, and jet fuel markups respectively, on the log of the RFS credit price in 2013 and 2014, while controlling for market concentration, market share, crude input prices, productivity, and the suite of fixed effects defined above. Recall from section 4 that in 2013, policy uncertainty and transportation infrastructure limitations caused a large spike in the RFS credit price with subsequent volatility carried through to 2014. I find that during this period, increases in the RFS credit price actually *increased* markups for conventional and reformulated gasoline, had no effect on markups for diesel fuel (testing conventional diesel and ultra-low sulfur diesel separately does not change the results), and *decreased* markups for jet fuel, a non-regulated product. The results are consistent with the pass-through results presented in the previous section. Columns 2 and 3 shows that in 2013-2014, a 10% (0.7 ¢/gallon) increase in the RFS credit price resulted in a 0.36% and a 0.25% increase in conventional and reformulated gasoline markups respectively. On average across both fuel types, this translates to a 0.33% or 0.1 ¢/gallon) increase in gasoline markups, or roughly an additional $2.9 million in additional revenue per month.

The results for jet fuel markups and marginal costs are surprising, particularly because jet fuel is not regulated under the RFS. Jet fuel is a refinery's third most valuable and highest volume product. As I will show in the following section, the production of jet fuel increased in response to increases in the RFS credit price. If demand for jet fuel was constant during this period, then changes in jet fuel production resulted in a shift in the jet fuel supply curve causing lower jet fuel prices and correspondingly lower markups.

Production Decisions and the RFS Credit Price

In a final application, I evaluate the effect of changes in the RFS credit price on the mix of refinery outputs. To do so, I regress the log of the product share for a given product on the log of the RFS credit price. Let the product share for product j produced by refinery i at time t be $PS_{ijt} = \frac{Q_{ijt}}{\sum_j Q_{ijt}}$, where Q_{ijt} is the quantity of product j for j=$\{$Gasoline, Diesel, Aviation Fuel, Jet Fuel$\}$, produced by refinery i at time t.[46] The regression is the following:

$$\ln PS_{ijt} = \gamma_0^j + \gamma_1^j \ln RFS_t + \gamma_2^j \boldsymbol{X}_{it} + \sigma_{iy} + S_s + \nu_{ijt}, \tag{21}$$

where $\ln RFS_t$ is the log of the RFS credit price and $\boldsymbol{X}_{it}$ includes the quality of crude oil such as API gravity and sulfur content as lower quality crude oil will produce more lower quality products, all else equal. Standard errors are clustered at the refinery-month level as the credit price varies at the monthly level while the errors may be correlated within refineries.

The results, presented in Table 12 provide evidence that refineries adjusted their outputs based on changes in the RFS credit price in 2013 and 2014. Increases in the RFS

[46] I limit the product shares to be out of the total production of gasoline, diesel, aviation fuel, and jet fuel because these are the most profitable products. Using the full sample of data yields similarly statistically significant results.

credit price throughout this period caused a decrease in the product shares of ultra low sulfur diesel (column 4) but caused an increase in the product share of jet fuel. Interestingly, the RFS credit price does not appear to have impacted the production of any other fuel.

Table 12: Production Decisions and the RFS Credit Price (2013-2014)

	CG	RFG	Diesel	ULSD	Avgas	Jet
	(1)	(2)	(3)	(4)	(5)	(6)
Log RFS Credit Price	0.023	0.012	0.045	-0.039**	-0.063	0.128***
	(0.019)	(0.043)	(0.070)	(0.016)	(0.234)	(0.036)
Firm-Year FE	Y	Y	Y	Y	Y	Y
Seasonal FE	Y	Y	Y	Y	Y	Y
Controls	Y	Y	Y	Y	Y	Y
R-squared	0.909	0.896	0.894	0.804	0.564	0.859
N	2399	894	1127	2169	168	2137

Notes: The dependent variables are the product shares of conventional gasoline, reformulated gasoline, regular diesel, ultra-low-sulfur diesel, aviation fuel, and jet fuel respectively. Controls include refinery level crude oil quality (API gravity and sulfur content). Standard errors clustered at the refinery-month level.
***Significance at the 1 percent level.
**Significance at the 5 percent level.
*Significance at the 10 percent level.

Using the observed product shares, the coefficients can be interpreted as follows: during 2013-2014, a 10% ($0.007/gallon) increase in the RFS credit price resulted in a 0.221 percentage point increase in the product share of jet fuel and 0.113 percentage point decrease in the product share of ultra low sulfur diesel. For example, the product share of jet fuel would have increased from 17% to 17.221%.[47]

This product substitution also has important environmental implications. For example, these results show that in 2013 a 10% increase in the RFS credit price would have resulted in an unintended increase in jet fuel production of approximately 301 billion gallons.[48] Using social cost of carbon estimates from the Environmental Protection Agency, I find that the additional production and subsequent consumption of jet fuel associated with a 10% increase in the RFS credit price would have resulted in additional emissions costs between $35-$179 million in 2013 alone.[49]

To provide some context for the magnitude of the emissions leakage, I relate the leaked emissions damage estimates to the avoided emissions damages under the RFS. As a consequence of the 2013 RFS mandate, roughly 13 billion gallons of ethanol and 1.4 billion gallons of biodiesel were consumed in the U.S. in place of conventional gasoline

[47]Between 2013-2014, the average product shares for ultra low sulfur diesel and jet fuel were 29% and 17% respectively (out of gasoline, diesel, aviation, and jet fuel). Therefore, a 1.3% increase in the product share of jet fuel would be translated into a 0.221 percentage point increase in the product share of jet fuel ($17\% * \gamma_1^{Jet} = 17\% * 0.013 = 0.221$), for example.

[48]Between mid-2012 and mid-2013, the RFS credit price increased by roughly 11.5 cents per gallon or 460%.

[49]The average quantity of jet fuel produced by refineries and blenders in the U.S. in 2013 was 1,510,192 barrels per day. The average amount of CO_2 produced from burning a gallon of jet fuel is 21.1 pounds per gallon. A 10% increase in the RFS credit price resulted in a 0.221 percentage point increase in the product share of jet fuel. Finally, the EPA's social cost of carbon numbers are $12, $40, and $62 per tonne of CO_2 for 5%, 3%, and 2.5% discount rates respectively.
Sources: http://www.eia.gov/environment/emissions/co2_vol_mass.cfm
http://www3.epa.gov/climatechange/EPAactivities/economics/scc.html

and diesel. I find that the emissions leakage associated with a 10% increase in the RFS credit price equates to roughly 3-5% of the avoided emissions associated with the RFS, depending on whether life-cycle or direct ethanol emissions are considered.[50] Although the damages from emissions leakage do not appear to completely undermine the effectiveness of the RFS, it is important to consider that, to date, it is unclear whether the life cycle emissions damages from ethanol are less than the emissions damages from conventional gasoline consumption. Ultimately, these results highlight the importance of accounting for the production process and the ability of firms to substitute away from regulated products when setting environmental regulations.

10 Incidence

In the following section I use the markup and pass-through results to estimate the incidence of uncertainty in the RFS credit price. Weyl and Fabinger (2013) define a general welfare incidence formula, which encompasses a range of market structures as

$$WI = \frac{\rho}{1-(1-\tilde{\theta})\rho}. \tag{22}$$

where $\tilde{\theta} = \left(\frac{p-\hat{mc}}{p}\right)\epsilon_d$ represents the conjectural variations parameter as a function of the Lerner index and the elasticity of demand.[51] When markets are perfectly competitive, the burden of a tax is split between producers and consumers and depends on the ratio of the supply and demand elasticities. In contrast, the burden of a tax is more than fully shared by producers and consumers when markets are imperfectly competitive because the equilibrium quantity is already lower than the efficient quantity. In other words, in a monopoly setting, the burden of a tax is greater than the revenue it raises as taxing a monopoly increases the existing deadweight loss associated with market power.

Increasing taxes in imperfectly competitive markets implies an additional social trade-off: market power increases producer surplus but creates deadweight loss (DWL). Weyl and Fabinger (2013) show that increased competition through exogenous entry, or small changes in output, reduces DWL and excess producer surplus in the same manner as changes in the tax rate. To capture this trade-off, Weyl and Fabinger (2013) define social incidence as the ratio between the change in DWL and the change in producer surplus relative to a change in an exogenous quantity entering the market $\tilde{q}$, $SI = \frac{dDWL/d\tilde{q}}{dPS/d\tilde{q}}$. Thus, lower measures of social incidence - producer surplus changes faster than DWL with a change in tax rates - imply greater societal benefit of taxation under imperfect competition. The social incidence can be calculated using the same three parameters from above via the following expression

[50] Emissions estimates for biofuels come from the California Air and Resources Board: http://www.afdc.energy.gov/data/10330

[51] The conjectural variations parameter is usually denoted simply as θ, however, I used θ to represent output elasticities previously in the paper. I therefore denote the conjectural variations parameter with a tilde.

$$SI = \frac{\tilde{\theta}\rho}{1+(1-\tilde{\theta})\rho}.$$

I estimate the welfare incidence and social incidence resulting from changes in the RFS credit price using the pass-through estimates reported in Table 10 and wholesale gasoline and diesel demand elasticities reported by Foster, Haltiwanger, and Syverson (2008) and Sweeney (2015). I estimate the Lerner index using the estimates of marginal costs from Section 8. The results are reported in Table 13. I report the welfare incidence and social incidence assuming monopoly, $\tilde{\theta} = 1$, perfect competition, $\tilde{\theta} = 0$, and the estimated conjectural variations parameter values, $\tilde{\theta} = \frac{p-\hat{mc}}{p}\epsilon_d$.

Table 13: Incidence and Welfare Changes By Fuel Type

	Gasoline (1)	Diesel (2)	ULSD (3)	Jet (4)
	Welfare Incidence $\frac{dCS}{dRFS}/\frac{dPS}{dRFS}$			
Estimated $\tilde{\theta} = \bar{\mu}\epsilon_d$	16.11	1.28	1.09	-.45
PC $\tilde{\theta} = 0$	-3.8	5.62	3.22	-.47
Monopoly $\tilde{\theta} = 1$	1.36	0.85	0.76	-0.9
	Social Incidence $\frac{dDWL}{d\tilde{q}}/\frac{dPS}{d\tilde{q}}$			
Estimated $\tilde{\theta} = \bar{\mu}\epsilon_d$	0.23	0.38	0.36	1.96
PC $\tilde{\theta} = 0$	0	0	0	0
Monopoly $\tilde{\theta} = 1$	1.36	0.85	0.76	-0.9

Notes: $\tilde{\theta}$ represents the conjectural variations parameter, estimated via the mean markup for each product, $\bar{\mu}$, multiplied by the wholesale demand elasticity for each product, ϵ_d. Alternatively, $\tilde{\theta}$ can be set to zero to simulate perfectly competitive conduct or set to 1 to simulate monopoly conduct. $dRFS$ refers to the change in the RFS credit price.

The incidence estimates produce several remarkable findings. First, I find that the burden of the unexpected shocks to the RFS credit price in the gasoline market is borne 16 times more by consumers than producers. Social incidence in the gasoline market is 0.23, implying producers bear the majority of the burden relative to the societal burden captured by DWL. In percentage terms, this implies the consumer share of the welfare incidence is 94% while the DWL share of the social incidence is only 19%. Second, I find that consumers also bear the majority of the burden of the RFS credit price in the diesel markets. I estimate welfare incidence in the diesel market to be 1.28 (consumer share is 56%) and in the ultra-low-sulfur diesel market to be 1.09 (consumer share is 52%). Third, I evaluate the incidence of the RFS on non-regulated fuel markets. Unsurprisingly, I find that the social incidence in the jet fuel market is 1.96, indicating that the change in DWL greatly outweighs the change in producer surplus. This result arises from the fact that the RFS caused an increase in jet fuel production and a decrease in jet fuel prices, reducing the DWL associated with market power in the jet fuel market, but not completely eliminating the excess producer profits.

When pass-through is less than 1 and greater than 0, the monopoly and perfectly competitive incidence estimates place upper and lower bounds on incidence. However, if pass-through exceeds 1, as in the gasoline market, the perfectly competitive incidence measure becomes negative indicating that either producer or consumer surplus is increasing in the RFS credit price. Intuitively, if the costs are more than fully passed onto consumers, producers are likely benefiting. Similarly, welfare incidence in the jet fuel market is always negative because pass-through in the jet fuel market is negative. This indicates that consumer surplus increased while producer surplus decreased in the jet fuel market.

The incidence results highlight the magnitude of the consequences associated with failing to account for market power when setting regulation. In particular, consumers bear the majority of the burden of the policy. In line with previous work on the Theory of the Second Best, these findings suggest multiple policy instruments are needed to internalize multiple market failures.

11 Conclusion

This paper makes three main contributions to two strands of literature. I estimate the impact of an incredibly important and understudied regulation, the Renewable Fuel Standard, on one of the largest and most complex industries in the U.S., the wholesale petroleum product market. First, I modify a novel production function methodology to estimate market power in the U.S. oil refining industry. The advantage of the methodology is that it does not require assumptions about demand curves, the nature of competition in the market, or market structure and allows me to estimate markups and marginal costs at particular points along the production schedule. The disadvantage of the methodology is that it does not allow for counterfactual simulations. The two key assumptions of the model are that firms minimize production costs and that input allocations are observed.

I then use the estimated markups and marginal costs to evaluate the impact of exogenous shocks in the RFS credit price on petroleum product prices, markups, marginal costs, and production decisions. I find that in 2013 and 2014, changes in the RFS credit price were more than fully passed onto wholesale gasoline prices. This finding stands in contrast to a large literature that finds less than or nearly perfect pass-through in many contexts. The result can be attributed to unique demand conditions and imperfect competition in the wholesale petroleum product market. Decomposing the pass-through rate shows that a 10% increase in the RFS credit price increased gasoline and diesel marginal costs by 0.08% and 0.18% respectively, but also increased gasoline markups by .33% on average. This implies that the costs of the regulation were excessively passed onto consumers and actually increased market power in the short run. In a similar vein, I find that increases in the RFS credit price caused firms to substitute non-regulated jet fuel production for regulated ultra-low-sulfur diesel production leading to an additional $35-$179 million in leaked emissions damages per year. Correspondingly, I find that jet fuel prices and markups decreased in response to increases in the RFS credit price, consistent with an outward shift in the supply curve for jet fuel.

Combined, the results in this paper provide empirical evidence on how the Theory of the Second Best plays out in an important and complex industrial setting. I show that 94% of the burden of short-run shocks in the RFS credit price were borne by consumers in the gasoline market. Consequently, uncertainty in the RFS credit price likely exacerbated existing welfare losses due to market power, by actually increasing regulated and non-regulated fuel markups. Moreover, I show that incomplete regulations in multi-product production settings allow firms to substitute regulated production for non-regulated production, which can attenuate the overall effectiveness of the policy.

Recovering markups and marginal costs requires the estimation of a refinery-product level production function. To do so, I use the Ackerberg, Caves, and Frazer (2015) proxy method to address unobserved refinery productivity and effort, and I perform robustness checks to provide confidence in my results. A beneficial feature of my dataset and the petroleum industry is that I can plausibly observe input allocation in a multi-product setting, which allows me to directly estimate a multi-product production function.

A limitation of the methodology used in the paper is that it is not fully structural in the sense that I do not recover demand and supply functions. I therefore am unable to estimate welfare effects or perform counterfactual policy experiments. For instance, the RFS is intended to have long-run dynamic implications for investment in renewable fuel production technology. It is unclear whether short-term shocks in the RFS credit price are causing refineries to invest in blending technology or to bank credits. If so, such investments will have future implications for the RFS credit market. Likewise, investment in cellulosic production technology will reduce the costs of biofuels, which will have an effect on the RFS credit price. Whether or not the RFS is having an impact on biofuel investment is an open question. More broadly, estimating a structural model to estimate welfare would allow one to weigh in on the overall efficiency of the RFS.

While this paper provides evidence of some peculiar effects of the RFS, future work might explore the mechanisms behind the pass-through and markup findings. For example, the RFS is only one of many policies that currently impact the petroleum product industry. I find greater than 100% pass-through in only two years of data while variation in the RFS credit price has continued through 2015 and 2016. One could assess the impact of similar policy shocks in other years, or estimate long run tax, marginal cost, and crude oil price pass-through to understand if the results are unique to 2013 and 2014 or unique to the RFS credit price. In a similar vein, Borenstein and Shepard (2002) show that firms pass-through increases in crude oil prices faster than similar decreases. Evaluating the pass-through of increases and decreases in the RFS credit price, looking at regional variation in pass-through, or looking at how observable refinery characteristics, such as vertical integration, impact pass-through could shed light on refinery behavior and competition. Finally, conditional on data acquisition, the analysis should be extended to retail price pass-though and should be separated by fuel type, i.e., E10 and E85.

References

Ackerberg, Daniel et al. (2007). "Econometric tools for analyzing market outcomes". *Handbook of Econometrics* 6, pp. 4171–4276.

Ackerberg, Daniel A, Kevin Caves, and Garth Frazer (2015). "Identification properties of recent production function estimators". *Econometrica* 83.6, pp. 2411–2451.

Auffhammer, Maximilian and Ryan Kellogg (2011). "Clearing the air? The effects of gasoline content regulation on air quality". *The American Economic Review*, pp. 2687–2722.

Berman, Eli and Linda TM Bui (2001). "Environmental regulation and productivity: evidence from oil refineries". *Review of Economics and Statistics* 83.3, pp. 498–510.

Borenstein, Severin, James B. Bushnell, and Frank A. Wolak (2002). "Measuring Market Inefficiencies in California's Restructured Wholesale Electricity Market". *American Economic Review* 92.5, pp. 1376–1405.

Borenstein, Severin and Andrea Shepard (2002). "Sticky prices, inventories, and market power in wholesale gasoline markets". *RAND Journal of Economics* 33.1, pp. 116–139.

Brown, Jennifer et al. (2008). "Reformulating competition? Gasoline content regulation and wholesale gasoline prices". *Journal of Environmental Economics and Management* 55.1, pp. 1–19.

Buchanan, James M. "External Diseconomies, Corrective Taxes, and Market Structure". *American Economic Review* 59.1, pp. 174–177.

Burkholder, Dallas (2015). "A Preliminary Assessment of RIN Market Dynamics, RIN Prices, and Their Effects". *Office of Transportation and Air Quality, US EPA, at http://www. regulations. gov.*

C.F.R. (2015). *Code Federal of Regulations, Title 40, Chapter 1, Subchapter C, Part 80, Subpart M, Section 80.1407.*

Collard-Wexler, Allan and Jan De Loecker (2015). "Reallocation and Technology: Evidence from the US Steel Industry". *American Economic Review* 105.1, pp. 131–71.

Cowing, Thomas G and V Kerry Smith (1977). "A note on the variability of the replacement investment capital stock ratio". *The Review of Economics and Statistics*, pp. 238–243.

De Loecker, Jan (2011). "Product differentiation, multiproduct firms, and estimating the impact of trade liberalization on productivity". *Econometrica*, pp. 1407–1451.

De Loecker, Jan and Pinelopi Goldberg (2014). "Firm Performance in a Global Market". *Annu. Rev. Econ* 6, pp. 201–27.

De Loecker, Jan and Frederic Warzynski (2012). "Markups and Firm-Level Export Status". *The American Economic Review* 102.6, p. 2437.

De Loecker, Jan et al. (2016). "Prices, Markups, and Trade Reform". *Econometrica* 84.2, pp. 445–510.

EIA (2014a). "Refinery Capacity Report". *online publication.*

— (2014b). "When was the last refinery built in the United States?" *online publication.*

EPA (2015a). "Renewable Fuel Standard". *online publication.*

Fabra, Natalia and Mar Reguant (2014). "Pass-Through of Emissions Costs in Electricity Markets". *American Economic Review* 104.9, pp. 2872–99.

Foster, Lucia, John Haltiwanger, and Chad Syverson (2008). "Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability?" *American Economic Review* 98.1, pp. 394–425.

Fowlie, Meredith L (2009). "Incomplete environmental regulation, imperfect competition, and emissions leakage". *American Economic Journal: Economic Policy* 1.2, pp. 72–112.

Gary, James H, Glenn E Handwerk, and Mark J Kaiser (2007). *Petroleum refining: technology and economics.* CRC press.

Goldberg, Pinelopi Koujianou and Rebecca Hellerstein (2008). "A Structural Approach to Explaining Incomplete Exchange-Rate Pass-Through and Pricing-to-Market". *American Economic Review* 98.2, pp. 423–29.

Hall, Robert E (1988). "The Relation between Price and Marginal Cost in US Industry". *The Journal of Political Economy*, pp. 921–947.

Irwin, Scott (2014). "Rolling Back the Write Down of the Renewable Mandate for 2014: The RINs Market Rings the Bell Again". *farmdoc daily* 4.148.

Knittel, Christopher R, Ben S Meiselman, and James H Stock (2015). *The Pass-Through of RIN Prices to Wholesale and Retail Fuels under the Renewable Fuel Standard.* Tech. rep. National Bureau of Economic Research.

Lade, Gabriel E, C-Y Cynthia Lin, and Aaron Smith (2015). *Policy Shocks and Market-Based Regulations: Evidence from the Renewable Fuel Standard.* Tech. rep. Working Paper.

Levinsohn, James and Amil Petrin (2003). "Estimating production functions using inputs to control for unobservables". *The Review of Economic Studies* 70.2, pp. 317–341.

Lipsey, R. G. and Kelvin Lancaster (1956). "The General Theory of Second Best". English. *The Review of Economic Studies* 24.1, pp. 11–32.

Muehlegger, Erich (2006). "Gasoline price spikes and regional gasoline content regulation: A structural approach".

Olley, G. Steven and Ariel Pakes (1996). "The Dynamics of Productivity in the Telecommunications Equipment Industry". English. *Econometrica* 64.6, pp. 1263–1297.

Seade, J (1985). *Profitable Cost Increases and the Shifting of Taxation: Equilibrium Response of Markets in Oligopoly.* Tech. rep. University of Warwick, Department of Economics.

Sweeney, Richard L. (2015). "Environmental Regulation, Imperfect Competition and Market Spillovers: The Impact of the 1990 Clean Air Act Amendments on the US Oil Refining Industry". *Job Market Paper.*

Weyl, E Glen and Michal Fabinger (2013). "Pass-through as an economic tool: Principles of incidence under imperfect competition". *Journal of Political Economy* 121.3, pp. 528–583.

Appendices

Appendix A: Bootstrap Routine

The bootstrap procedure is as follows.

Step 1: Estimate the production function (10). The coefficient estimates from the original dataset are called β^o for original.

Step 2: Draw a random sample of refinery observations, with replacement, from the observed sample of refineries. This means taking a refinery's full set of observations in each draw. Do this until the original number of observations is reached.

Step 3: Re-estimate (10) and keep the new coefficient estimates. Call these estimates $\beta^{b,n}$ for bootstrapped.

Step 4: Repeat steps 2 and 3 n times. For the current draft of this paper, $n = 20$.

Step 5: Compute the mean output elasticities from expression (17) for each set of coefficient estimates and take the standard deviation of the set of mean output elasticity estimates. The output elasticity point estimates are the mean estimates from the original dataset, i.e., $\theta^o_{ijt}(\beta^o)$.

Step 6: Compute a set of markups for each bootstrapped dataset including the vectors of output elasticities $\theta^{b,n}_{ijt}(\beta^{b,n})$ and $\theta^o_{ijt}(\beta^o)$. Markup standard deviations can be computed from the $n + o$ markup estimates.

Step 7: Estimate equation (20) for each set of markups. Then compute standard errors for the coefficient estimates based on the bootstrapped iterations of (20).