Essays in Industrial Organization

By

Ke Zhang

A DISSERTATION SUBMITTED IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY
(ECONOMICS)

at the

UNIVERSITY OF WISCONSIN – MADISON

2012

Date of final oral examination: June 29, 2012

The dissertation is approved by the following members of the Final Oral Committee:

Jean-Francois Houde, Assistant Professor, Economics
Kenneth Hendricks, Professor, Economics
Alan Sorensen, Associate Professor, Economics
Amit Gandhi, AssistantProfessor, Economics
Brent Hueth, Associate Professor, Agricultural & Applied Economics

Abstract

Chapter 1 Impact of the Sulfur Content Regulation on Refinery Input Choices

The U.S. Environmental Protection Agency decided in 2000 to enact a policy change that reduced the amount of sulfur permitted in diesel fuel. The goal of this paper is to examine how the policy change affected the decision process of a refinery in choosing the optimal type of crude oil to be used as raw material. This paper uses panel data on U.S. refineries from 1994 to 2009 to determine how did refinery choices in the type of crude oil used change with the implementation of the new policy.

Chapter 2 Estimation of a Production Function with Differentiated Input

Unlike traditional heterogeneity in total factor productivity, refineries produce a mix of outputs and each refinery has a comparative advantage for a different mix of products. In order to accurately capture the production of petroleum outputs for economic analysis, refinery technology and crude oil characteristics must be taken into account. To model the multiple outputs, I use a CES production function to capture the substitutability between each output. In this paper I generalize the share parameter by allowing the output share to be a function of both plant and crude oil characteristics.

Chapter 3 The Effect of Sulfur Content Regulation on the U.S. Diesel Market

The goal of this paper is to quantify the effects of the policy change on the diesel fuel market. I will investigate the impact on the marginal cost of diesel and other petroleum products for U.S. refineries. Also, I will quantify the effects on the diesel wholesale prices in each state and determine how each region in the U.S. may be impacted differently by the new regulation.

Acknowledgements

I first want to thank Jean-Francois Houde for his support and advice while I was working on my dissertation. JF were always patient and provided honest and insightful input on my work. I also want to thank Ken Hendricks, Alan Sorenson and Amit Gandhi for their efforts and suggestions for my research and during my job market process.

I would also like to express my appreciation for the friends I've met during my years of graduate school. In particular I want to thank Hillary Caruthers, Camila Casas, Tim Heugerich, Tim Lee, Rebecca Lessem, Rui Li, Peter Newberry, Michael Rapp, Carl Sanders, Jiao Shi, Yuya Takahashi and Xiao Wang.

Finally I want to thank my family for their encouragement and support.

Contents

Abstract **i**

Acknowledgements **ii**

1 Impact of the Sulfur Content Regulation on Refinery Input Choices **1**

1.1 Introduction 1

1.2 Background 4

1.3 Data 6

1.3.1 Refinery Characteristics 6

1.3.2 Crude Oil Characteristics 7

1.4 Methodology 9

1.4.1 Difference in Difference 9

1.4.2 Regresion Discontinuity 10

1.5 Regression Analysis 11

1.6 Conclusion 14

1.7 Appendix 15

2 Estimation of a Production Function with Differentiated Input **26**

2.1 Introduction 26

2.2 Model 28

2.3 Data 35

2.3.1 Refinery Characteristics 35

2.3.2 Crude Oil Characteristics 35
2.3.3 Crude Oil Price 36
2.3.4 Other Prices 39
2.4 Estimation 41
2.5 Result 44
2.5.1 Parameter Estimates 44
2.5.2 Goodness of Fit 46
2.5.3 Marginal cost & Profit Margin 47
2.6 Conclusion 47
2.7 Appendix 49

3 The Effect of Sulfur Content Regulation on the U.S. Diesel Market 55
3.1 Introduction 55
3.2 Environmental Policy 57
3.3 Methdology 59
3.3.1 Demand 59
3.3.2 Refinery Objective 60
3.3.3 Market Equilibrium 62
3.4 Result 63
3.5 Conclusion 66
3.6 Appendix 67

Bibliography 73

Chapter 1

Impact of the Sulfur Content Regulation on Refinery Input Choices

1.1 Introduction

The United States is the largest consumer of crude oil in the world, consuming approximately 19 million barrels of crude oil each day, around 23% of the world's total production. These crude oil make up 35% of America's energy needs with much of that goes into transportation. In order to reduce pollutants from vehicle emission, the U.S. Environmental Protection Agency decided in 2000 to enact a policy change that reduced the amount of sulfur permitted in diesel fuel. The policy was implemented in 2006 and it limits the maximum sulfur content in 80% of diesel outputs from large refineries to 15 parts per million. By 2010, the restriction was applied to all diesel outputs from U.S. refineries. Since all crude oil contains various amounts of sulfur, the policy change increased the burden of U.S. refineries to remove the sulfur from crude oil. The goal of this chapter is to examine how the policy change affected the decision process of a refinery in choosing the type of crude oil to be used as raw material. I estimate the effect of the policy using both regression discontinuity and difference in difference to examine how does the plant's choices vary with production technologies.

In 2000, the EPA announced a comprehensive national control program to regulate heavy duty vehicle and its fuel as a single system. As a part of its program it aims to reduce the level sulfur in highway diesel fuel by 95%. There are two reasons for the reduction of sulfur in high way diesel fuel. Firstly sulfur emission from motor vehicles is a major contributor to acid rain, which causes both serious health and economical impacts across the U.S. Secondly, the sulfur released also damages the newly required catalytic emission control devices in vehicles which are used to reduce nitrogen oxides and other fine particles been released into the air.

Prior to the regulation change, the maximum sulfur level in diesel fuel was 500ppm, also known as Low Diesel Fuel (LDF). The program restricted refiners to start producing 80% of their diesel fuel with a sulfur content of no more than 15ppm, Ultra Low Diesel Fuel (ULDF), by June 1st of 2006. By June 2010, full conversion to ULDF was required for all refiners. Outline of the restrictions are in table 20. The program also provided hardship provisions to two types of refineries. It delayed the onset of the restriction to small refineries and Geographic Phase-in (GPA) Refineries. Small refiners are defined as refineries with less than 1500 employees and less than 155,000 barrels per day crude processing capacity and GPA refineries are refineries in the Rocky Mountains[1] and Alaska. Small refiners and GPA refiners were permitted to continue to produce LDF till 2010, by which they must also switch over to ULDF.

---

[1] States included are New mexico, Utah, Colorado, Wyoming, Idaho, Montana and North Dakota

As far as I am aware, there hasn't been no literature that attempt to investigate the effect of sulfur content regulation in diesel fuel. However, there has been a large number of literature that investigates the effect of the Clean Air Act policy on oxygenated gasoline and refomulated gasoline. Most recently Brown, Hastings Mansur and Villas-boas[4] and Chouinard and Perloff[7] investigate the effect of the gasoline policy on prices and how the effect vary with geographic segmentation and market concentration. Auffhammer and Kellog[2] examines the effect of the gasoline content policy on air quality.

Refineries produce a mix of outputs and each refinery has a comparative advantage for a different mix of products. Some plants are relatively more productive in producing lighter petroleum products such as gasoline, while other plants are more productive in producing heavier products such as diesel fuel. The mix of outputs by a refinery depends on two factors. One is the difference in production technology across plants, the other is the difference in the type of crude oil available to each refinery. Both processing technology and the type of crude oil used for production dictate the mix of outputs that is achievable. This chapter uses panel data on U.S. refineries from 1994 to 2009 to determine how did refinery choices in the type of crude oil used change with the implementation of the new policy. I will also examine how does the choice in the type of crude oil relate to the production technology of a refinery.

1.2 Background

Crude oil is the main input used in the production of petroleum products and the differentiation in the type of crude oil is an important aspect of the production process. The two main characteristics of crude oil are API Gravity and sulfur content. API Gravity measures the average density of the crude oil and is a key measure that determines the output distribution of the crude oil. Higher API Gravity implies lighter crude oil. **??** shows the API Gravity of various petroleum products and how crude oil with different densities can impact the mix of outputs. Propane is the lightest with an API Gravity of around 105 and fuel oil, which is used in power plants or ship boilers, is the heaviest with an API Gravity of around 10. A barrel of crude is a mixture that contains both light oil and heavy oil. Crude oil is referred to as light if it contains a higher percentage of light oil than heavy oil. Therefore, lighter crude oil contains more oil that is similar to gasoline compared to a heavier crude oil. Table **??** shows how as oil gets heavier, the percentage of heavier oil it contains increases and therefore it results in more heavy outputs such as residual fuel. Since different refineries have different levels of technology and the mix of outputs may vary, each refinery will have its preferred type of crude oil. Refineries that produce more heavier outputs will prefer heavier crudes and refineries that produce more lighter outputs will prefer light crude.

Crude oil is also vertically differentiated by the amount of sulfur it contains. All crude oil contains various amounts of sulfur with some crudes contain as much as 4% sulfur. Also heavier crude oil generally contains more amounts of sulfur. There are two main reasons why sulfur needs to be removed. One is that sulfur is very corrosive and damages equipment. The other is that sulfur content is limited by federal regulations. Therefore crude oil that contains more sulfur is more costly for refineries since more sulfur has to be removed.

Once crude oil is separated according the density in a process known as distillation, each refinery has a variety of technologies at its disposal to alter the crude oil. Diagram 1 provides an illustration of these processes. These technologies after distillation are known as downstream processes and can be divided into 3 major categories of technology. The first category includes reforming, isomerisation and alkylation, and is typically used to upgrade light crude oil such as butane into gasoline. They are illustrated as upgrading technology on the diagram. The other major component, described as conversion technology, includes coking, catalytic cracking and hydrocracking. Their purpose is into convert heavy crude oil, which are used in boilers, into lighter products such as gasoline and diesel fuel. These two categories of technology are summarized in table 2. I measure each group of technologies by aggregating the capacity of each components within the group. Both upgrading capacity and conversion capacity are expressed as a percentage of the capacity of processes under the two categories relative to the overall capacity of the plant. The overall capacity of the plant is the total amount of crude oil that can be processed each day and is typically measured using the distillation capacity.

Table 3 shows how conversion technology changes the output mix of products. Using the same type of crude oil, either light crude or heavy crude, conversion technology decreases the percentage of output in the heavier categories such as residual fuel and increases the output of lighter products such as gasoline. Increasing the capacity in conversion technology allows even greater shifts from heavy petroleum products into lighter petroleum products. The last category in the refining process is the removal of sulfur. The purpose of removing sulfur is to reduce sulfur dioxide (SO_2) emission from vehicles, power plants and etc. After distillation, sulfur compounds can be found in all output streams, however it is much more prevalent in the heavier outputs such as diesel and residual fuel. Currently EPA regulates that the level of sulfur in diesel fuel to be no more than 15 parts per million (ppm). Prior to the policy change, the maximum amount of sulfur in diesel was 500ppm.

1.3 Data

1.3.1 Refinery Characteristics

For plant characteristics, I use data provided by the Refinery Capacity Report, an annual survey conducted by the Energy Information Administration (EIA). The survey is conducted for all refineries located in U.S. states and territories and response is mandatory for all refineries. The data is available from 1994 till 2009, except no survey was conducted in 1996 and 1998. The survey includes the capacity of all major refinery processes.[2] I aggregate the capacity of each individual process into the three main categories, upgrading, converting and desulfur. Table 4 shows the summary statistics of the capacity of the three main technology categories.

---

[2] Methodology of the survey is available at www.eia.gov/petroleum/refinerycapacity/820notes.pdf

The table shows the overall processing capacity and the capacity of different technologies for all U.S. refineries from 1994 to 2009 except the two years missing during that period. The overall capacity is in thousands of barrels per day and the statistics for each technology is the percentage the input capacity of these technologies in terms of the overall input capacity. It shows the maximum possible capacity for these technologies to be in excess of 100%. The reason is because some refineries process the input through the same technology multiple times. The summary statistics also shows that there are large variations in the capacity of these technologies across U.S. refineries. Diagram 2 shows the the number of plants over the period and the average refinery capacity in plants that existed at both the beginning and end of the data period. It shows there has been an almost 30% increase in plant size during the 15 year period. However if plants that are shutdown or restarted are included, the increase is average capacity will be higher. Diagram 3 shows the changes increase in the capacity of the 3 main types of technologies across U.S. over the same time period. The capacity in conversion and upgrading has remained fairly stable over the past 15 years, however the capacity in sulfur removal increased by 25% after the policy was introduced in 2000.

1.3.2 Crude Oil Characteristics

For the characteristics of inputs used by each refinery, I use the monthly Company Level Imports dataset from the EIA. The dataset is a shipment level record of all crude oil imported into the U.S. and it provides the API gravity and sulfur percentage of each shipment. The dataset also records the port of arrival, company name and the city name of processing facility. Since the shipment data records at the company level, but the model estimates the input choices at refinery level, each shipment needs to be assigned to each individual refinery. The shipments are assigned based on the processing facility's city and company name as recorded in the dataset. In majority of the cases, approximately 67% of all shipments, the processing facility's city name matches the refinery's location and the company owns no other refineries in the Metropolitan Statistical Area (MSA). If the import facility's city is unknown, then they are matched using port of entry of the shipment and refineries that are located along oil pipelines that can connect the refinery to the port of entry.

Using map of crude oil pipelines as illustrated in diagram 4, I track all the refineries that are owned by the importing firm and located along the pipeline from the port of entry. Then the shipment is divided between these refineries located along the pipeline using the overall capacity of the refinery as weights. If there is no pipeline connecting the port of entry and any refineries owned by the importing company, the shipment is assigned to all the refineries owned by the importing company in the same PADD[3].

The other scenario is if a firm has multiple facilities in a region and the import data only shows shipments to a strict subset of these facilities. For example, for several years, Suncor owned two refineries in the Denver area, west Denver and Commerce City on the east side. Shell also operated two refineries in the same region for a period, Wilmington and Bakersfield. The import shipment data, however,only lists Denver or

---

[3] The U.S. is divided into 6 Petroleum Administration for Defense Districts. They are as follows, East Coast, Midwest, Gulf, Rocky Mountains, West Coast and Puerto Rico.

Los Angeles for the shipments instead of the specific refinery location. In this case, I will assumed the shipment is divided equally between the two refineries weighted by the overall capacity of each plant. Using these methods, 83% of the shipments are assigned to a refinery. The remaining unassigned shipments are shipments to National Petroleum Reserve and energy trading companies. Summary statistics of the assigned shipments and total shipments are reported in table 6. The first four columns report the summary statistics for assigned shipments and the remaining four columns record the statistics for all shipments. Table 7 reports the standard deviation of both all assignment shipments and shipments to East coast and the Gulf. The standard deviation between plants is twice the standard deviation within a plant. This shows persistence in the type of crude oil used by a plant which correspondes to the story that each plant has an optimal type of crude oil that maximizes profit.

1.4 Methodology

1.4.1 Difference in Difference

In the DD approach, the identification of the effect of the sulfur cotent regulation relies on the year to year differences in the type of crude oil after controlling for the annual shifts. I estimate the plant level crude oil characteristics on plant technology and policy change. The estimating equation is given as,

$$y_{it} = \alpha_{1i} + \delta_t + \beta_1 k_{it} + \beta_2 D_{it} + \beta_3 D_{it} k_{it} + \epsilon_{it}. \tag{1.1}$$

y_{it} are the two characteristics of crude oil used by plants, API density and sulfur level. α_{1i} controls for plant level fixed effects. δ_t are controls for yearly and monthly fixed effects. The yearly fixed effect controls for any shifters that causes aggregate changes in the type of crude oil used across periods. The monthly fixed effect capture the year-to-year differences in crude oil characteristics. k_{it} are plant technology variables that impact the type of crude oil used. D_{it} is the dummy variable for the policy change. For the control group, I use refineries located in the Rocky Mountains who are excluded from the initial policy application. I use 18 months prior and after the implementation of the policy as the sample size of the estimation.

1.4.2 Regresion Discontinuity

In the RD approach, identification of the effect of the sulfur content regulation relies on the narrow window of time just prior and after the implementation of the policy. I vary the size of the window between 6 months and 18 months before and after the implementation. I also vary the control group, I use Rocky Mountain refineries and small refineries as the control group separately and together. The estimating equation is

$$y_{it} = \alpha_i + \beta_1 k_{it} + \beta_2 D_{it} + \beta_3 D_{it} k_{it} + g(t) + \epsilon_{it}. \tag{1.2}$$

y_{it} are the two plant level crude oil characteristics used by each refinery. α_i is the firm fixed effect. k_{it} is a vector of plant characteristics and D_{it} is the dummy for the policy decision. $g(t)$ is a function of time where June 2012 is the point of discontinuity. For the choice of $g(t)$, I use a polynomial form and fixed effect in separate estimations.

1.5 Regression Analysis

If the type of crude oil used by a refinery is dependent on the technology of the plant, then the data will exhibit persistence where the type of crude oil used by a refinery. In table 9, I regress the two characteristics of the crude oil used by each refinery on the cruded oil used in last period. Column 1 shows the result for sulfur and column 4 shows the result for API Gravity. The result indicates strong persistence in the crude oil used by a refinery. Refineries that tend to purchase heavy sour crude oil in previous periods will most likely purchase similar crude oil in the next period. In columns 2 and 5, I add characteristics of plant technology and regional fixed effect. The coffeficient for the regression of sulfur on desulfur indicates that plant that have more desulfur technology tend to purchase more sour crude oil. The result also shows that the type of crude oil purchased is correlated with the location of the plant. Plants that are on the West coast (PADD 4) and East coast (PADD 1) generally use crude oil that contains less sulfur than refineries located in other regions.

The result is similar for API Gravity. It shows that plants that have higher capacity for conversion and desulfur tend to use heavier crude oil and plants that have higher capacity to upgrade tend to use lighter crude oil. It also shows that refineries located on the West Coast generally use lighter crude oil than other regions in the U.S. In both regressions, even after controlling for plant characteristics and regional fixed effect, the result still shows persistence in the typde of crude oil across periods. After controlling for plant fixed effect in the regressions reported in column 3 and 6, the correlation between periods is much lower. This indicates plant level unobserved heterogeneity that affects the type of crude oil purchases by each plant.

Table 10 shows regression results from estimating the effect of the policy on the density of crude oil using regression discontinuity. The first column shows the result using the time window 6 months prior to and after the policy implementation. It also uses a polynomial form for the time variable. Column two uses the same time window, however a more general fixed effect for time is used instead of a polynomial form. Column three to column five extends the time frame to 18 months prior to and after the policy implementation. It also uses a fixed effect for time. Column three isolates the control group to small refineries and column four isolates the control group to refineries to Rocky Mountains. Column five shows the result from the regression that allows interaction between the policy and plant technology characteristics. Table **??** shows the regression results on the effect of the policy on the sulfur level of the crude oil used by refineries. Table 12 shows regression results from the difference in difference approach. In column 1, the dependent variable is the density of crude oil and in column 2, the dependent variable is the sulfur level.

All the results show that higher capacity in upgrade technology is correlated with lower density crude oil with less sulfur content. Plants that have capacity in desulfur and conversion is correlated with denser crude oil with higher sulfur content. This relationship is expected. Upgrade technology is used to process light crude oil to make high octane gasoline, therefore the positive correlation between high upgrade capacity and lighter crude oil with less sulfur is expected. Desulfur and conversion technology is used process heavy crude oil and break them up into lighter components. Therefore, it also correspondes with the positive correlation from the regression.

The result also shows that the policy had a statistically significant impact on the type of crude oil used by refineries. As a result of the policy, the result shows that plants increase the API density of crude oil by 0.8 to 1.3 units. This is approximately a 2.5% to 4% decrease in the density of crude oil. It also shows that the sulfur level decreases by approximately 0.2 percentage points. This represents a 10% reduction in the sulfur level in crude oil.

The effect of the policy also varies with the technology of refineries. The estimation shows that refineries with higher capacity in upgrade technology decreases the density of crude oil more than refineries with lower capacity in upgrade technology. For conversion technology, the estimation result varies with the two approaches. In RD, it shows that refineries with high conversion capacity purchases heavier crude oil after the policy change. In DD, the result is not statistically significant. In either case, it shows that refineries became more specialized after the policy change. Refineries that specialize more in high quality gasoline using upgrade technology become even more specialized by shifting to lighter crude oil. On the other hand, the estimation does not show that the decrease in sulfur level vary with plant technology. However, this may be due to the increase in desulfur capacity from 2003 to 2006. During that period refineries with desulfur technology at the top 25% at the beginning increased their capacity by 12% on average, where as refineries with desulfur technology at the bottom 25% increased their capacity by 60% on average.

1.6 Conclusion

In this chapter, I examined how the policy implementation is correlated with a change in the use of crude oil by refineries. The result shows that after the policy change, refineries generally used crude oil that are both lighter and contain less sulfur. The regression result also shows that the change in choices after the policy implementation varied across regions. We observe the greatest change in refineries located on the East Coast, whereas refineries in Midwest saw little changes in the type of crude oil used. The regression result also shows that the choice of crude oil is correlated with the capacity in the technology of refineries. Refineries that have higher capacity in upgrade technology generally purchase lighter crude oil. Refineries that have higher capacity in conversion and desulfur technology generally purchase heavier crude oil with more sulfur. Since changes in the type of crude oil used in production may affect the mix of outputs produced by a refinery, therefore in order to investigate the effect of the environmental policy on the petroleumn products market, it is important to control for the effect of the policy on the choice of crude oil used in production.

1.7 Appendix

Table 1: Petroleum Products and types of Crude Oil

Product	API Gravity	Light Crude	Heavy Crude
Propane/Butane	105	3%	1%
Gasoline	55	30%	14%
Distillate (Jet Fuel, Home Heating Oil, Diesel Fuel)	45	34%	22%
Residual Fuel	10	33%	63%

Sweet Crude are crude oil with API Gravity above 34 API. Heavy Crude are crude oil with API Gravity below 24. Source: Valero

Table 2: Production Process under each Technology Category

Upgrade	Conversion
Reforming	Catalytic Cracker
Isomerisation	Hydrocracker
Alklyation	Coker

Table 3: Petroleum Product Outputs under different levels of Conversion Technology

Technology	Product	Light Crude	Heavy Crude
No Conversion	Propane/Butane	3%	1%
	Gasoline	30%	14%
	Jet Fuel/Diesel	34%	22%
	Residual Fuel	33%	63%
High Conversion	Propane/Butane	7%	6%
	Gasoline	58%	44%
	Jet Fuel/Diesel	28%	32%
	Residual Fuel	15%	26%

Light Crude are $>$ 34 API. Heavy Crude are $<$ 24 API. Source: Valero

Table 4: Summary Statistics on Refinery Technology

All Periods				
Process	Mean	Std.Dev.	Min	Max
Capacity	121	113	1	596
Upgrade	0.24	0.18	0	1.29
Conversion	0.44	0.32	0	1.51
Desulfurization	0.60	0.29	0	2.22

Capacity in thousands of barrel per day. Others are percentages.

Table 5: Summary statistics on utilization rate

	Mean	Std.Dev	Min	Max
Overall	91.16	8.18	43	109.4
Between		4.39		
Within		7.02		

Table 6: Summary Statistics on crude oil imports

	Cleaned Data				Original			
Process	Obs	Mean	Min	Max	Obs	Mean	Min	Max
Quantity	103248	430	1	10672	123869	436	1	10672
Sulfur Level	103240	1.55	1.27	8	123869	1.55	0.01	8
API Gravity	103246	29.58	0.24	82.6	123869	29.58	0.24	82.6

The unit for quantity is thousands of barrels. The unit for Sulfur Level is %.

Table 7: Standard deviation for crude oil imports

		Std.Dev	Between	Within
All Sample	API Gravity	6.74	6.99	3.23
	Sulfur	1.10	1.07	0.53
East Coast & Gulf	API Gravity	7.27	7.29	2.89
	Sulfur	1.14	1.15	0.49

Table 8: Summary statistics on source of imports

	Mean	Std.Dev	Min	Max
Importing countries by each plant	4.45	1.77	1	14

Table 9: Regression - Lag Effect

	Sulfur	Sulfur	Sulfur	API Gravity	API Gravity	API Gravity
Upgrade		-0.026	-0.415		0.408	0.754
		(0.027)	(0.110)		(0.179)	(0.790)
Conversion		0.034	0.469		-0.434	-1.561
		(0.019)	(0.098)		(0.128)	(0.618)
Desulfur		0.087	0.324		-0.546	-4.831
		(0.016)	(0.088)		(0.101)	(0.540)
Lagged	0.910	0.697	0.267	0.896	0.685	0.230
	(0.004)	(0.004)	(0.012)	(0.004)	(0.004)	(0.012)
Padd 2		0.021			0.071	
		(0.009)			(0.092)	
Padd 3		0.035			0.153	
		(0.013)			(0.086)	
Padd 4		0.039			-1.310	
		(0.014)			(0.109)	
Padd 5		-0.087			1.229	
		(0.015)			(0.099)	
Padd 6		-0.043			0.171	
		(0.035)			(0.229)	
Constant	0.139	0.059	0.859	3.058	3.880	22.341
	(0.007)	(0.015)	(0.050)	(0.119)	(0.169)	(0.533)
Fixed Effect			Firm			Firm

Table 10: RD - Dependent Variable API Gravity

	API Gravity	API Gravity	API Gravity	API Gravity	API Gravity
Upgrade	17.704*	17.763*	11.360*	8.591*	7.227*
	(1.211)	(1.217)	(0.813)	(0.578)	(0.848)
Conversion	-7.403*	-7.406*	-7.332*	-9.845*	-5.365*
	(1.022)	(1.026)	(0.621)	(0.379)	(0.681)
Desulfur	-3.609*	-3.650*	-3.440*	-5.182*	-2.991*
	(0.749)	(0.753)	(0.462)	(0.312)	(0.461)
Policy	0.763*	0.878*	1.427*	1.307*	0.331
	(0.347)	(0.370)	(0.337)	(0.232)	(0.860)
Time	3.133				
	(16.448)				
Time Sqrd	-0.003				
	(0.015)				
Policy*Upgrade					14.852*
					(1.668)
Policy*Conversion					-5.384*
					(1.234)
Policy*Desulfur					-0.065
					(0.956)
Time Fixed Effect		yes	yes	yes	yes
Time Smaple	6 Mth	6 Mth	18 Mth	18 Mth	18 Mth
Control Group	Both	Both	Small Ref.	Rocky Mt.	Both
Firm Fixed Effect	yes	yes	yes		yes
Adjusted R^2	0.46	0.46	0.39	0.38	0.41

Table 11: RD - Dependent Variable Sulfur Content

	Sulfur	Sulfur	Sulfur	Sulfur	Sulfur
Upgrade	-1.513*	-0.777*	-0.763*	-1.372*	-0.308*
	(0.209)	(0.210)	(0.138)	(0.102)	(0.144)
Conversion	0.085	0.085	0.315*	0.895*	0.222*
	(0.176)	(0.177)	(0.105)	(0.067)	(0.116)
Desulfur	0.477*	0.479*	0.603*	1.127*	0.590*
	(0.129)	(0.130)	(0.078)	(0.055)	(0.078)
Policy	-0.164*	-0.156*	-0.144*	-0.251*	-0.465*
	(0.077)	(0.081)	(0.057)	(0.117)	(0.146)
Time	3.753*				
	(2.835)				
Time Sqrd	-0.003				
	(0.003)				
Policy*Upgrade					-0.231
					(0.283)
Policy*Conversion					0.168
					(0.269)
Policy*Desulfur					-0.005
					(0.162)
Time Fixed Effect		yes	yes	yes	yes
Time Smaple	6 Mth	6 Mth	18 Mth	18 Mt	18 Mth
Control Group	Both	Both	Small Ref.	Rocky Mt.	Both
Firm Fixed Effect	yes	yes	yes		yes
Adjusted R^2	0.41	0.41	0.38	0.37	0.43

Table 12: Difference in Difference

	API Gravity	Sulfur
Upgrade	15.905*	-1.935*
	(1.585)	(0.278)
Conversion	-14.110*	0.815*
	(1.049)	(0.184)
Desulfur	-0.236	0.426*
	(0.737)	(0.129)
Policy	0.816*	-0.165*
	(.282)	(0.075)
Policy*Upgrade	4.904*	-0.265
	(1.862)	(0.326)
Policy*Conversion	1.647	-0.271
	(1.281)	(0.224)
Policy*Desulfur	-1.429	-0.093
	(0.970)	(0.170)
Time Fixed Effect	yes	yes
Time Smaple	18 Months	18 Months
Control Group	Rocky Mountains	Rocky Mountains
Plant Fixed Effect	yes	yes
Adjusted R^2	0.49	0.49

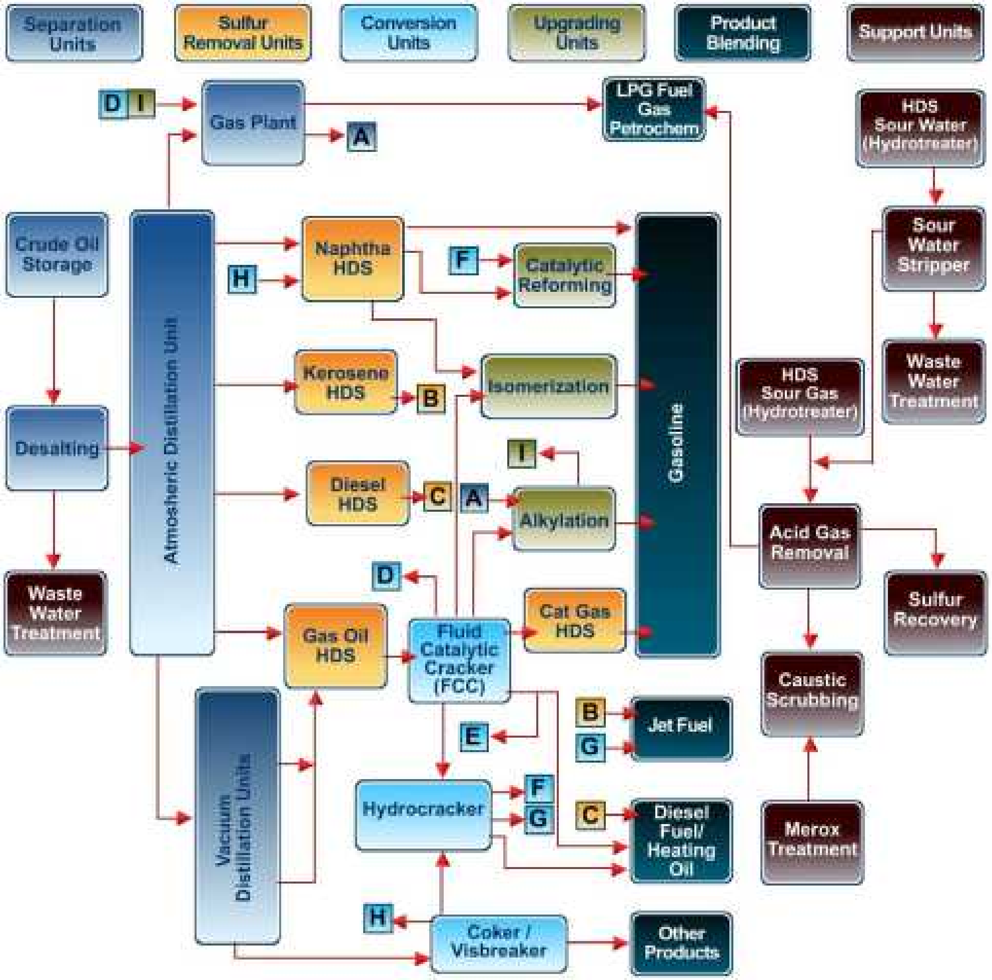

Figure 1: Diagram of Refinery Technology. Source: Veolia Water

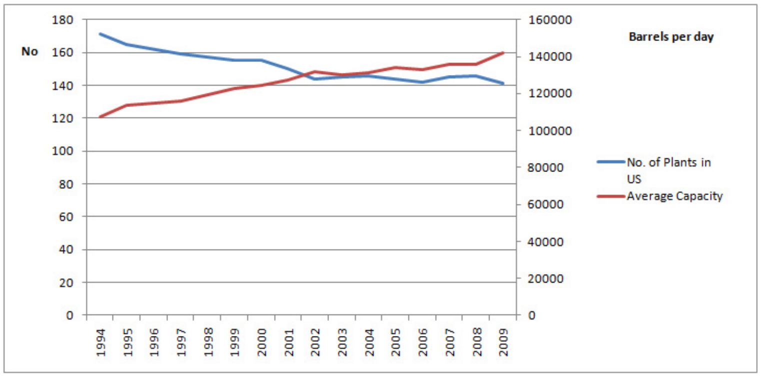

Figure 2: Number of refineries in U.S. and average capacity for plants that appeared during the whole period.

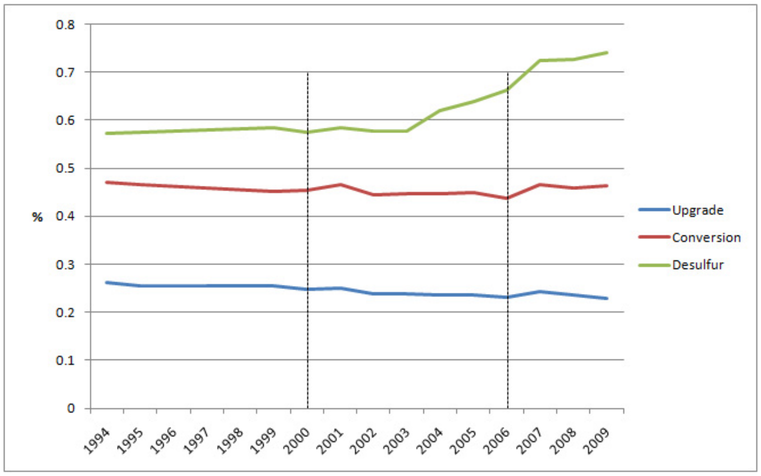

Figure 3: Capacity of three cateogies of technologies for plants that appeared during the whole period.

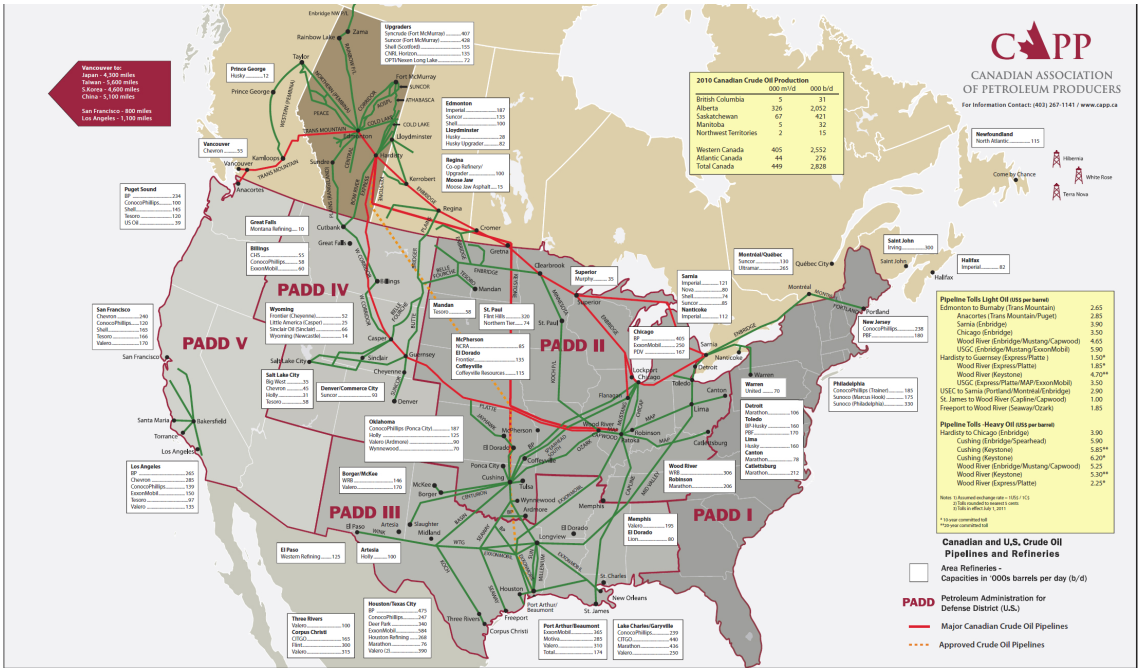

Figure 4: Map of crude oil pipelines

Chapter 2

Estimation of a Production Function with Differentiated Input

2.1 Introduction

Unlike traditional heterogeneity in total factor productivity, refineries produce a mix of outputs and each refinery has a comparative advantage for a different mix of products. Some plants are relatively more productive in producing lighter petroleum products such as gasoline, while other plants are more productive in producing heavier products such as diesel fuel. The two main factors that determine the comparative advantage of a refinery are plant technology and the characteristics of the crude oil used. Both processing technology and the type of crude oil used for production dictate the mix of outputs that is achievable. In order to accurately capture the production of petroleum outputs for economic analysis, refinery technology and crude oil characteristics must be taken into account. To model the multiple outputs, I use a CES production function to capture the substitutability between each output. The CES production function was first introduced in Arrow, Chenery, Minhas and Solow (1961) [1] for the single output and multiple input scenario. Vinod (1968) [16] and Mundlak (1963) [14] adopted it for the multiple output scenario. In traditional CES, the output share, which determines the mix of outputs, is treated as a constant. In this chapter I generalize this assumption by allowing the output share to be a function of both plant and crude oil characteristics. Therefore I am able to capture how variations in plant technology and crude oil differentiation give rise to different production possibility frontier for each plant. I estimate the generalized multi-output production function using panel data on U.S. east coast and gulf coast refineries from 1994 to 2009 with GMM.

As far as I am aware, there have only been two papers that have attempted to follow a structural approach to model the petroleum products production process. There are typically two assumptions made with regard to the production technology of refinery industry in these papers. One is that the distribution of petroleum products is fixed across plants and over time such as Chesnes (2009) [6]. However data shows the distribution of output is not fixed. For example, in December 2010, output of gasoline was 70% of diesel production along the Texas gulf coast, whereas only 3 years earlier, Texas gulf coast was producing 60% more gasoline than diesel fuel. The ratio between gasoline and diesel fuel is neither fixed across plants. For December of 2010, even though Texas gulf coast was producing more diesel fuel than gasoline, U.S. east coast was producing 8 times more gasoline than diesel fuel. The second assumption, which is used by Muehlegger (2006)[11], is that the main input, crude oil, is homogeneous and the petroleum products are perfect substitutes as outputs. It assumes that a barrel of crude oil can be separated into any combination of outputs without any efficiency loss and it does not have an effect on the cost of production. These papers also do not take into account that refineries differ in their technologies.

2.2 Model

In order for each plant to choose the optimal output mix to maximize its profit, I use a multiproduct production function with constant elasticity of substitution (CES) on the output side. The production function assumes the quantity of outputs and quantity of inputs are separable, implying that the quantity of input does not affect the ratio of outputs, but only the overall quantity of outputs.

Each plant is indexed by i. This is a static model where the time subscript t is removed for readability. The unit of time for each time period is one month. The capacity for the technologies of a plant: upgrading, converting and desulfur are denoted by k_{1i}, k_{2i} and k_{3i} respectively and the capital K_i denotes the vector of the technologies $\{k_{1i}, k_{2i}, k_{3i}\}$. All three variables are measured as a percentage of the capacity of crude oil that each technology can process relative to the overall number of barrels a plant can process. The overall capacity of the plant is denoted as k_{0i}. Each output product is indexed by j and the quantity, in number of barrels per day, produced for each product j is denoted as q_{ji}. There are five outputs: propane, gasoline, jet fuel, diesel and residual fuel. The API Gravity[1] of the crude oil is denoted as a_i and it is normalized to be between zero and one. The sulfur content percentage is denoted as s_i. The total quantity of crude oil used in barrels per day is z_i. The production function is given as

$$\begin{aligned} &\textstyle\sum_{j=1}^{J}(B_j(a_i, K_i)q_{ji}^{\rho})^{\frac{1}{\rho}}A_i = z_i \\ &s.t. \quad \textstyle\sum_j B_j(a_i, K_i) = 1 \;\; \forall a_i \in [0,1], s_i \in [0,1] \\ &s.t. \quad B_j(a_i, K_i) > 0 \;\; \forall j, a_i \in [0,1], s_i \in [0,1] \end{aligned} \tag{2.1}$$

---

[1] Higher API Gravity implies lighter crude oil

The left hand side of the equation is the output component. The Total Factor Productivity (TFP) shock is denoted by A_i. The elasticity of substitution is captured through the variable ρ. The share parameter is captured through the function $B_j(a_i, K_i)$. The share parameter for each product is conditional on the technologies of the plant and the API gravity of crude oil. With the inputs fixed, the left hand term generates a concave production frontier curve with a constant elasticity of substitution. It behaves in a similar manner compared to the traditional CES production function, except the generated curve is concave rather than convex. As ρ approaches positive infinity, the output mix approaches fixed ratio similar to a Leontieff production function. As ρ approaches 1, the production production approaches the perfect substitutes form with a linear frontier curve and if $\rho = 2$, then the elasticity of substitution is equal to 1. The production frontier curve created by this production function form is illustrated in diagram 5.

$B_j(a_i, K_i)$ shifts the slope of the frontier curve by altering the productivity for each output. The productivity for product j is decreasing on B_j. A higher B_j implies q_{ji} requires more quantities of inputs. The parameter is a function of two sets of variables, characteristics of the crude oil used, which is the API Gravity a_i and plant characteristics K_i. The model allows the API Gravity of the crude oil, a_i, to affect the share parameter of each product, B_j, differently. If B_j is increasing on a_i, which is expected for heavier products such as diesel fuel, then it implies heavier crude oil will increase the output for that particular product. On the other hand B_j is expected to be decreasing on a_i for lighter products, implying heavier crude decreases the output of lighter products such as propane. The term a_i changes the ratio $\frac{B_1(a_i,K_i)q_{1i}^{\rho-1}}{B_2(a_i,K_i)q_{2i}^{\rho-1}}$, thereby altering the marginal rate of technical substitution. So, in the case where the density of the crude oil decreases, then $\frac{B_{gas}(a_i,K_i)q_{gasi}^{\rho-1}}{B_{diesel}(a_i,K_i)q_{dieseli}^{\rho-1}}$ is expected decrease, causing the plant to increase the production of gasoline relative to diesel. This is illustrated in diagram 6. The green line illustrates the production possibility frontier curve using heavier crude oil and the red line illustrates the curve using lighter crude oil.

Alternatively, a plant can achieve similar results as a change in its input type by having different capacities in their technology, $K_i = \{k_{1i}, k_{2i}\}$. The model assumes a substitutive relationship between the density of crude oil and plant conversion capacity. Instead of purchasing lighter crude, a firm can increase its capacity in conversion technology, k_2, and convert heavier crude oil into lighter crude oil. The conversion capacity, k_{1i}, change the ratio in $\frac{B_{gas}(a_i,K_i)q_{ji}^{\rho-1}}{B_{diesel}(a_i,K_i)q_{ji}^{\rho-1}}$ and achieve the same result as shown in diagram 6. At each time period, the technology of a plant is taken as given, however it will affect the type of crude oil used. Compared to a plant with higher conversion capacity, low conversion plants have to use lighter crude oil in order to produce the same mix of outputs.

There are also two constraints on the model. Firstly, the model assumes

$$\sum_j B_j(a_i, K_i) = 1 \tag{2.2}$$

for any choices of a_i and K_i. This is the standard normalization assumption in traditional CES production function. The other assumption is that $\rho > 1$. This ensures that the production possibility frontier is concave.

The functional form chosen for the share parameter function is

$$B_j(a_i, K_i) = \frac{\exp(D_j(a_i, K_i))}{1 + \sum_{l=2}^{5} \exp(D_l(a_i, K_i))} \; j \neq 1 \tag{2.3}$$

$$D_j(a_i, K_i) = \alpha_{j1} + \alpha_{j2}a_i + \alpha_{j3}k_{1i} + \alpha_{j4}k_{2i} + \alpha_{j5}a_ik_{1i} + \alpha_{j6}a_ik_{2i} \; j \neq 1 \tag{2.4}$$

The share parameter is dependent on the API gravity of the crude oil a_i, capacity of upgrade technology k_{1i} and the capacity of conversion technology k_{2i}. It follows a logistic function form because it allows the function $B_j(a_i, K_i)$ to aggregate to one and captures the relative productivity between outputs. Since the share function measures the relative productivity between products, in order for this model to be identified, one of the products must be assumed to be the outside good. In this case, the outside good is assumed to be propane and it follows the function form,

$$B_1(a_i, K_i) = \frac{1}{1 + \sum_{l=2}^{5} \exp(D_l(a_i, K_i))}. \tag{2.5}$$

Thus, if $D_j(a_i, K_i)$ is decreasing, it implies that the productivity of product j is increasing relative to the productivity of propane. The first term of the marginal rate of technical substitution between product j and propane is $\frac{B_j(a_i,K_i)}{B_1(a_i,K_i)} \frac{q_{ji}^{\rho-1}}{q_{1i}^{\rho-1}} = D_j(a_i, K_i) \frac{q_{ji}^{\rho-1}}{q_{1i}^{\rho-1}}$ and it will decrease with the change in $D_j(a_i, K_i)$.

For the parameters in the share function, α_j measures the impact of the characteristics on the product share. α_{j2} captures the effect of the density of crude oil on the productivity on each product j. If an output prefers lighter crude oil compared to other outputs, then the α_{j1} of this output will be smaller than the coefficients of other products. Similar, α_{j3} and α_{j4} reflects the effect of capacity in upgrade technology and cracking technology on the productivity of various outputs. α_{j5} and α_{j6} measures the effect due to the interaction between the type of the crude oil and the technologies of the refinery.

Other than crude oil, the other main input used in processing is natural gas. Natural gas is used as a fuel for processing crude oil and remove sulfur. The sulfur content of the crude oil is denoted as s_i. The maximum sulfur level under EPA regulation for each product is denoted as $\bar{s}_j$. The capacity of sulfur removal is denoted as k_{3i}. Also the density of product j is denoted as $\bar{a}_j$. The quantity of natural gas required to process the crude oil and to remove the sulfur so that the federal standard $\bar{s}_j$ are is given by the function

$$NG_i = \sum_j w(a_i, \bar{a}_j, s_i, \bar{s}_j, k_{3i}) q_{ji} \tag{2.6}$$

The function $w(a_i, \bar{a}_j, s_i, \bar{s}_j, k_{3i})$ captures how much natural gas is required to process the crude oil in order to meet the standards of product j. Then the natural gas required for each product is aggregated weighted by the quantity of each output. There are two processes that depend on natural gas. Firstly, natural gas is used as fuel to break up crude oil from the density level a_i to the density $\bar{a}_j$. $\bar{a}_j$ is the approximate density of product j. Secondly, it is used to decrease the sulfur level from s_i to $\bar{s}_j$. The level $\bar{s}_j$ is determined by the government regulation. There are also two technology shocks, $\epsilon = \{\epsilon_1, \epsilon_2\}$, that affect the quantity of natural gas required. One of the technology shocks captures the technology heterogeneity in the breaking up of crude oil and enters multiplicatively with API Gravity, a_i. The other technology shock captures the technology heterogeneity in sulfur removal and enters multiplicatively with the sulfur content level. $w(a_i, \bar{a}_j, s_i, \bar{s}_j, k_{3i})$ is in the form of

$$w(a_i, \bar{a}_j, s_i, \bar{s}_j, k_{3i}) = \epsilon_1(\bar{a}_j - a_i))_+^{\lambda_1} k_{i2}^{\lambda_2} + \epsilon_2(s_i - \bar{s}_j)_+^{\lambda_3} k_{i3}^{\lambda_4} \tag{2.7}$$

The component $\epsilon_1(\bar{a}_j - a_i))_+^{\lambda_1} k_{i2}^{\lambda_2}$ captures the amount of natural gases needed to break up the crude oil. Since higher a_i implies lighter oil, the difference between the required density of the product and the density of the crude oil is denoted by $(\bar{a}_j - a_i)_+$. $\epsilon_2(s_i - \bar{s}_j)_+^{\lambda_3} k_{i3}^{\lambda_4}$ captures the amount of natural gas needed to remove the sulfur to the required federal standard. The excess sulfur above the federal standard is captured through $(s_i - \bar{s}_l)_+$.

Each plant chooses its distribution of output and type of input in two stages. In the initial stage, each plant chooses the cost minimizing type of inputs conditional on the output mix. I assume each plant take the input prices as given. The price of crude oil is given by a continuous function $P_l(a_i, s_i)$. It is a hedonic price where the price depends on the choice of input type by the refinery and the location of the refinery l. Also I use the capital $\mathbb{Q}_i$ to denote the vector of firm's outputs $\{q_{1i}, \ldots, q_{Ji}\}$ for readability. The price of natural gas is denoted by the variable P_l^{ng}. The plant also faces soft capacity limit in $\gamma_1(\frac{z_i}{k_{0i}})^{\gamma_2}$. The cost due to capacity will increase exponentially. Thus the plant's problem in the initial stage is

$$\begin{aligned} C(\mathbb{Q}_i) &= \min_{z_i,a_i,s_i,NG} P(a_i,s_i)z_i + P^{ng}NG_i + \gamma_1(\tfrac{z_i}{k_{0j}})^{\gamma_2} \\ s.t.\ \ & \textstyle\sum_{j=1}^{J}(B_j(a_i,K_i)q_{ji}^{\rho})^{\frac{1}{\rho}}A_i = z_i \\ & NG_i = \textstyle\sum_j w(a_i,\bar{a}_j,s_i,\bar{s}_j,k_{3i})q_{ji} \end{aligned} \tag{2.8}$$

Solving for the optimal $a^*(\mathbb{Q})$ and $s^*(\mathbb{Q})$, gives us the cost curve as a function of the outputs.

$$\begin{aligned} C(\mathbb{Q}_i) = \ \ & P(a_i^*(\mathbb{Q}_i),s_i^*\mathbb{Q}_i)(A_i\textstyle\sum_{j=1}^{J}B_j(a_i^*(\mathbb{Q}_i),K_i)q_j^{\rho})^{\frac{1}{\rho}} + P^{ng}\textstyle\sum_j w(a_i^*(\mathbb{Q}_i),\bar{a}_j,s_i^*(\mathbb{Q}_i),\bar{s}_j,k_{3i})q_j \\ & +\gamma_1(\tfrac{(A_i\sum_{j=1}^{J}B_j(a_i^*(\mathbb{Q}_i),K_i)q_j^{\rho})^{\frac{1}{\rho}}}{k_{0j}})^{\gamma_2} \end{aligned} \tag{2.9}$$

Given the cost function, each plant choose the profit maximizing quantities for each output. In addition, the output goods' market is also assumed to be perfectly competitive and each plant is a price taker with P_j as the price of product j. According to the Energy Information Administration of the DOE, there are at least 60 refinery firms across U.S. operating each year since 1994. Using the operating capacity of these plants, I calculate that the mean Herfindahl index across U.S. is approximately 0.05 from 1994 till 2009. The low market concentration and because petroleum products are homogeneous at the wholesale level[2], perfect competition does not seem overly restrictive. The plant's profit maximizing problem becomes

$$\begin{aligned} &\max_{\mathbb{Q}_i} \textstyle\sum_j p_j q_{ji} - C(\mathbb{Q}_i) \\ &s.t. \quad q_{ji} \geq 0 \ \ \forall j \end{aligned} \ . \tag{2.10}$$

---

[2] Brand specific additives are added to gasoline after wholesale

2.3 Data

2.3.1 Refinery Characteristics

For plant characteristics, I use data provided by the Refinery Capacity Report, an annual survey conducted by the Energy Information Administration (EIA). The survey is conducted for all refineries located in U.S. states and territories and response is mandatory for all refineries. The data is available from 1994 to 2009, except no survey was conducted in 1996 and 1998. The survey includes the capacity of all major refinery processes.[3] I aggregate the capacity of each individual process into the three main categories, upgrading, converting and desulfur.

2.3.2 Crude Oil Characteristics

For the characteristics of inputs used by each refinery, I use the monthly Company Level Imports dataset from the EIA. The dataset is a shipment level record of all crude oil imported into the U.S. and it provides the API gravity and sulfur percentage of each shipment. The dataset also records the port of arrival, company name and the city name of processing facility. Since the shipment data records at the company level, but the model estimates the input choices at refinery level, each shipment needs to be assigned to each individual refinery. The shipments are assigned based on the processing facility's city and company name as recorded in the dataset. In majority of the cases, approximately 67% of all shipments, the processing facility's city name matches the refinery's location and the company owns no other refineries in the Metropolitan Statistical Area

---

[3] Methodology of the survey is available at www.eia.gov/petroleum/refinerycapacity/820notes.pdf

(MSA). If the import facility's city is unknown, then they are matched using port of entry of the shipment and refineries that are located along oil pipelines that can connect the refinery to the port of entry.

Refineries typically imports from several countries in order to mix and match to obtain the desired crude oil type since each oil has a distinct characteristics for its crude oil. Imported crude oil only makes approximately 60% of all crude oil used in the U.S. and the percentage of the imported crude oil vary greatly depending on the location of the plant. Table 13 shows the percentage of imported crude oil that makes up the total input used from 1994 to 2009 conditional on PADD. However no data is available on the domestic crude oil used by each plant. Therefore the chapter restricts the estimation to only the east coast (PADD 1) and the gulf (PADD 3).

2.3.3 Crude Oil Price

In order to estimate firm's input choices, the model needs the cost of crude oil to be a continuous function on the crude oil's characteristics controlling for the location of the refinery, $P_l^{cost}(a, s)$. Two sets of data are used to construct the hedonic price equation, the F.O.B price of major world oil streams and the weighted average acquisition for U.S. refineries. Both sets of data are published by the EIA. The EIA publishes weekly average F.O.B. prices on 38 major world oil fields. Majority of the sources are covered from 1994 till now. For the characteristics of the each oil stream, I use the data from the 2006 Crude oil Handbook published by Energy Intelligence. Energy Intelligence is a private company specializing in information and data on global energy industry. The firm conducts analysis of all the major world oil fields and the most recent characteristics of these oil fields are reported in its Crude oil Handbook. Using the first data, I run a hedonic regression and obtain the F.O.B price conditional on the time, characteristics of the crude oil and the region of the oil field. I use the second data to recover the cost of transportation between the refinery and the oil field.

Each oil field is indexed by n and time is denoted as t. The observed FOB for oil field n is P_{nt}^{FOB} and I estimate hedonic regression,

$$P_{nt}^{FOB} = X_{nt} + \beta_1 a_n + \beta_2 s_n + \mu_{nt} \tag{2.11}$$

where X_{nt} controls for year, month, and the continent of the oil field. The linear form on the explaining variables simplifies it's use in the production function. A more general form can be used, however, it will greatly complicate the model. This regression gives $\hat{P}_t^{FOB}(a, s, reg)$, the predicted F.O.B. price of the crude oil conditional on the characteristics, time and region of the oil field. After I recover the hedonic price conditional on the characteristics of the crude oil, I calculate the transportation cost. To obtain the transportation cost, I use the import shipment data to calculate the predicted average FOB price for all U.S. crude oil imports and compare that with the reported average acquisition cost.

Let each crude oil shipment be indexed by m. Using $\hat{P}_t^{FOB}(a, s, reg)$, I can calculate the FOB price of each shipment $\hat{P}_t^{FOB}(a_m, s_m, reg_m)$ where the characteristics of the shipment is denoted by a_m and s_m and region of the exporting country be denoted as reg_m. Then each individual shipment level FOB price is aggregated with the quantity of each shipment, Q_m as weights, $\sum_m(\hat{P}_t^{FOB}(a_m, s_m, reg_m))Q_m$, to obtain the weighted average FOB price for each refinery region. The actual weighted average acquisition cost for refinery in location l as reported by the EIA is denoted as P_{lt}^{cost}. The distance of the shipments from the import data is denoted as $Dist_m$. I estimate the transportation cost conditional on the distance between the location of the refinery and oil field

$$P_{lt}^{acq.cost} - \frac{\sum_m(\hat{P}_t^{FOB}(a_m, s_m, reg_m))Q_m}{\sum_m Q_m} = \beta_3 \frac{\sum_m(Dist_m)Q_m}{\sum_m Q_m} + \mu_{lt}. \tag{2.12}$$

Using the estimated parameters, I can calculate the acquisition cost of crude oil for a refinery located in region l purchasing from oil field n,

$$\begin{aligned} \hat{P}_{lnt}^{cost}(a_n, s_n) = \;& \hat{P}_t^{FOB}(a, s, reg) + \beta_3 Dist_{ln} \\ & X_{nt} + \beta_1 a_n + \beta_2 s_n + \beta_3 Dist_{ln}. \end{aligned} \tag{2.13}$$

The previous euqations determines the expected acquistion cost for each refinery in state l from the oil field n. This allows a refinery to choose which oil field, denoted as n, to purchase crude from, giving the acquisition cost, and characteristics of that oil field, $\{\hat{P}_{lnt}^{cost}, a_n, s_n\}$. However under the current model specification, the decision of the refinery is the characteristics pair $\{a, s\}$, the density and sulfur level of the crude oil maximize it's profit. Therefore, in order to obtain an aquisition price for every $\{a, s\}$ pair, I mix and match several oil fields such that the average characteristics of the oil fields is equal to the $\{a, s\}$ pair, $\frac{\sum_n q_n a_n}{\sum_n q_n} = a$ and $\frac{\sum_n q_n s_n}{\sum_n q_n} = s$. The other objective in choosing the bundle of oil fields is to minimize the acquisition cost. The objective is

$$P_l^{acq.cost}(a, s) = min_{q_1,\ldots,q_n,\ldots,q_N} \hat{P}_{lnt}^{cost}(a_n, s_n) \tag{2.14}$$

such that $\frac{\sum_n q_n a_n}{\sum_n q_n} = a$ and $\frac{\sum_n q_n s_n}{\sum_n q_n} = s$

After the lower envelop is calculated for a grid of $\{a, s\}$ pairs, a smother function is calculated from the set of points. Figure 7 shows the hedonic price for the two characteristics of crude oil. It shows the change in price for every percentage change in API Gravity or sulfur content. The increase in the price of the two characteristics corresponds with the decrease in crude oil production at the beginning of the Gulf War and price returned to long term mean during the onset of the world financial crisis in 2008.

2.3.4 Other Prices

In addition, I use monthly prices for natural gas and various petroleum products published by the EIA. EIA provides several types of prices for natural gas, such as citygate, residential, commercial and industrial. The dataset for industrial prices does not cover all the years in the estimation, therefore citygate prices are used. Citygate prices measure the price at the station where distributing gas utility receives gas from a pipeline company. EIA also provides two types of prices for petroleum products, sale to end users and resale prices. Resale price is the wholesale price and is the priced faced by refiners. In addition, for gasoline, the average price across various octane level is used.

For petroleum products, each refinery sells each product to multiple number of states.

I observe the monthly output prices in each states, however, instead of solving a multi-market equilibrium for estimation, I assume each refinery faces a weighted average of the prices and the refinery determines the distribution of its output across regions is dependent only on the distance between the market and the refinery. The weight of the prices will depend on the market share of petroleum product sales to each region. I use map of petroleum pipelines to locate which states are potential markets for each refinery and using the pipelines I calculate the distance between the refinery located in state l and the potential market l', denoted by $dist_{ll'}$.

Then using data on the flow of petroleum products between PADDs in the US that is provided by the EIA, I calculate the share of outputs a refinery in state l produces that flows into each region l', denoted by $\frac{q_{ll'j}}{\sum_{l'} q_{ll'j}}$. Using the distance between the refinery and the market, and the percentage of a plant's output that flows into that market, I estimate the discount rate conditional on the distance.

$$\frac{q_{ll'j}}{\sum_{l'} q_{ll'j}} = \frac{exp(-\alpha_5 dist_{ll'})}{\sum_{l'} exp(-\alpha_5 dist_{ll'})}. \tag{2.15}$$

After obtaining α_5, the weight based on distance, I calculate the weighted price

$$p'_{lj} = \sum_{l'} \frac{exp(-\alpha_5 dist_{ll'}) P_{jl'}}{\sum_{l'} exp(-\alpha_5 dist_{ll'})} \tag{2.16}$$

faced by a plant in region l for product j.

2.4 Estimation

I estimate the production function using generalized method of moments. The estimated parameters are the coefficients in the output share $\{\alpha_{11}, \ldots, \alpha_{13}, \ldots, \alpha_{5,2}\}$ and $\{\theta_1, \ldots, \theta_5\}$, the elasticity of substitution ρ, coefficients for natural gas use $\{\lambda_1, \ldots, \lambda_6\}$, and coeffiecents for soft capacity constraint $\{\gamma_1, \gamma_2\}$. The estimation uses a combination of both macro and micro moments. Since I do not observe total quantity at the plant level, plant level total factor productivity shock A_i can not be obtained. However, since refining district level[4] production data is available, I assume that total factor productivity shock is common across plants in a region.

In order to generate the micro moments, I first solve first order conditions that are obatained from the objective function of each plant. The first order condition with respect to a_i is

$$\begin{aligned} a_i : \quad & \left(\tfrac{\partial P_l^c(a_i,s_i)}{\partial a_i}\right) z_i \\ & +(P_l^c(a_i, s_i) + \gamma_1\gamma_2 \tfrac{z_i^{\gamma_2 - 1}}{K_0^{\gamma_2}})\tfrac{1}{\rho} \textstyle\sum_j (B_j(a_i, K_i) q_{ji}^{\rho})^{\frac{1-\rho}{\rho}} A_i \textstyle\sum_j (\tfrac{\partial B_j(a_i,K_i)}{\partial a_i} q_{ji}^{\rho}) \\ & +P^{ng} \textstyle\sum_j \tfrac{\partial w_j(a_i,s_i,K_i)}{\partial a_i} q_{ji} = 0 \end{aligned} \tag{2.17}$$

where

$$\frac{\partial B_j(a_i, K_i)}{\partial a_i} = \begin{array}{l} \left[\tfrac{\exp(D_j(a_i,K_i))}{(1+\sum_{l\neq 1}\exp(D_l(a_i,K_i))}\right] \times \\ \left[\begin{array}{l} (\alpha_{j2} + \alpha_{j5} k_{1i} + \alpha_{j6} k_{2i})\times \\ \quad (1 + \sum_{l\neq j} \exp(D_l(a_i, K_i))) \\ -\sum_{l\neq j}(\alpha_{l2} + \alpha_{l5} k_{1i} + \alpha_{l6} k_{2i})\times \\ \quad \exp(D_l(a_i, K_i)) \end{array}\right] \end{array} \tag{2.18}$$

---

[4] Each refining district consists of several states. The east coast has two refining districts. The gulf has 5 refining districts.

$$\frac{\partial B_j(a_i,K_i)}{\partial a_i}=\frac{\left[\ -\sum_{l\neq 1}(\alpha_{l2}+\alpha_{l5}k_{1i}+\alpha_{l6}k_{2i})\times\exp(D_l(a_i,K_i))\ \right]}{(1+\sum_{l\neq 1}\exp(D_l(a_i,K_i)))^2} \tag{2.19}$$

and

$$\frac{\partial w_j(a_i,s_i,K_i)}{\partial a_i}=\epsilon_1\lambda_1(a_i-\bar{a}_j)^{\lambda_1-1}K_{crack}^{\lambda_1} \tag{2.20}$$

The first component is the effect of the weight on price of crude oil. The second component is the effect of the weight on the quantity of crude oil needed. The third component is the effect of the weight on the quantity of natural gas needed.

The first order condition with respect to s_i is

$$s_i:\quad \tfrac{\partial P_l^c(a_i,s_i)}{\partial s_i}z_i+P_l^{ng}\sum_j\tfrac{\partial w_j(a_i,s_i,K_i)}{\partial s_i}q_{ji}=0 \tag{2.21}$$

where

$$\frac{\partial w_j(a_i,s_i,K_i)}{\partial s_i}=\begin{cases}\epsilon_2\lambda_3(s_i-\bar{s}_j)^{\lambda_3-1}K_{sulfur}^{\lambda_4} & s_i>\bar{s}_j\\ 0 & otherwise\end{cases} \tag{2.22}$$

Through the two first order condition of the crude oil type on the cost function, they allow the two technology shocks to be identified. Parameters in the cost function are identified through variations in the price of natural gas and the aggregate quantity of natural gas used.

$$z_i:\quad z_i=\sum_j(B(a_i,K_i)q_{ji}^{\rho})^{\frac{1}{\rho}}A_i \tag{2.23}$$

$$\left(\frac{(p_{jl} - P_l^{ng} w_j(a_i, s_i, K_i))B_{-j}(a_i, K_i)}{(p_{-jl} - P_l^{ng} w_{-j}(a_i, s_i, K_i))B_j(a_i, K_i)} \right)^{\frac{1}{\rho - 1}} = \frac{q_{ji}}{q_{-ji}} \bigg|_{a_i = a_i^*, s_i = s_i^*} \tag{2.24}$$

The district level total factor productivity shock is obtained from aggregating the inputs and outputs for plants in each district.

$$A \sum_i \in l \sum_j (B_j(a_i, K_i) q_{ji}^{\rho})^{\frac{1}{\rho}} = z_l^* \tag{2.25}$$

where A is the common TFP shock across plants in a district and z_l^* is the observed total input for the refining district l.

Solving the FOCs, I can obtain the technology shocks ϵ_1 and ϵ_2. For the technology shocks, I allow for plant fixed effect and trend over time. The plant fixed effect will control for selection bias in the technology of the plant. The time trend controls for aggregate changes in the technology of plants.

$$\ln \epsilon_{it} = \delta_f + \mu Time + \eta_{it}, \tag{2.26}$$

where δ_f is the firm fixed effect. Both technology shocks η are assumed to equal to zero conditional on the instruments. The firm fixed effect is removed from $\ln \epsilon_{it}$ by removing the mean shock across plants within the firm. The moment conditions are $\frac{1}{IT} \sum_{it} \eta_{it} X_{it} = 0$, where X_{it} are the instruments. The instruments used for the micro moments are the price for each of the 5 products and the 3 observed plant characteristics

For the macro moments, the difference between the calculated aggregate product quantity and observed product quantiy $\sum_i q_{ji} - q_j^*$ is assumed to be conditionally equal to zero. The macro moments in this case is $\frac{1}{TL}\sum_{tl}(\sum_{i\in l} q_{tji} - q^{*}_{tlj})X_{tl} = 0$. q^{*}_{tlj} is the observed total output for product j in region l at time t and X_{tl} are the instruments. The other macro moment is the quantity of natural gas. Since the quantity of natural gas is observed at regional level, the calculated natural gas quantity is matched with the observed quantity, $\frac{1}{TL}\sum_{tl}(\sum_{i\in l} NG_{ti} - q^{*}_{tl})X_{tl} = 0$.The instruments used for the macro moments are the district level output prices, and district level average plant characteristics.

2.5 Result

2.5.1 Parameter Estimates

Table 14 reports the point estimates for the parameters in the production function. The first column is the effect of crude oil density on the productivity of various outputs. A positive coefficient implies the productivity of the product relative to propane is decreasing as the crude oil becomes lighter. In this case the result come out as expected. The coefficients for all the products are positive, this implies the produtivity of all these outputs relative to propane is decreasing as crude oil becomes lighter. The ordering of coefficients between the products imply residual fuel prefers heavier crude oil whereas gasoline is more productive with lighter crude oil. The second column is the effect of plant capacity in upgrade technology on the productivity of each output. The fourth column reports the interaction of upgrade technology and density. The result shows upgrade technology is important for the productivity of gasoline and jetfuel. The negative coefficient for gasoline gasoline implies higher capacity in upgrade technology implies higher productivity for gasoline and jetfuel over propane. The effect also becomes greater as crude oil becomes lighter. The third column is the impact of conversion technology on the relative productivity of each output. The results shows the coefficient for diesel and residual fuel being bigger than the coefficient of gasoline is statistically significant. This implies higher conversion capacity increases the relative productivity of gasoline compared to diesel and residual fuel. The negative coefficients on the interaction between API and conversion capacity implies the effect is greater when denser crude oil is used.

Results for the estimated parameters in the function for natural gas demand are listed in table 15. It shows that as increased sulfur in crude oil increases the marginal cost of production and as the amount of sulfur in crude oil increases, the marginal cost of removing that sulfur increases at an increasing rate. Table 16 shows the effect of capacity constraint on the marginal cost.

Figure 8 shows the simulated production possibility frontier using the estimated parameters. It compares two types of plants, low capacity conversion and high capacity conversion. The red curve is the production possibility frontier for a high conversion plant and the green curve is the production possibility frontier for a low conversion plant. The result shows that an increase in conversion technology increases the productivity of gasoline relative to diesel. Figure 8 shows the comparison between gasoline and propane and it compares the differences between low upgrade capacity and high upgrade capacity. The red curve is for a plant with low capacity in upgrade technology and the green curve is for a plant with high capacity in upgrade technology. It shows that plants with higher capacity in upgrade has relatively higher productivity in gasoline production compared to propane production. The result indicated by the diagram is expected since the purpose of the upgrade technology is to increased the productivity of gasoline relative to even lighter petroleum products as indicated in figure 9.

2.5.2 Goodness of Fit

Furthermore I test the fit of the model by simulating the weighted average output distribution of refineries across the east coast and comparing it to observed average. The result of the simulation is presented in table 17. For majority of the petroleum products, the model does a fairly good job in estimating the output percentage. However it slightly over-estimates output for propane, jetfuel and residual. The simulated output percentage over-estimates the actual output level by 40%. The most likely reason for the over-estimation of propane production is because the elasticity of substitution is assumed to be constant between outputs in the model. Since petroleum products can be ordered by their density level, it is strange to assume that the elasticity of substitution between propane and gasoline is the same as between propane and diesel, which, compared to gasoline, is much more heavier than propane. Therefore the model may have under-estimated the elasticity of substitution between propane and other products, resulting in over-estimating the production of natural gas. One method to address this issue is to use a nested CES production function so that the elasticity of substitution may vary depending on the output pair.

2.5.3 Marginal cost & Profit Margin

Using the estimated parameters, I am estimate the expected marginal cost on gasoline excluding crude oil cost on the technologies of plants. The result is reported in table 18. It shows that an increase in upgrade capacity by 10% lowers marginal cost by 1.15 dollars per barrel and an increase in conversion capacity by 10% lowers marginal cost by 1.83 dollars per barrel.

I also calculate the expected plant level profit margins per barrel and investigate how does the price margin vary with the technology of the plant. I run an OLS regression on the calculated profit margins with the upgrade capacity and cracking capacity as dependent variables. The result is in table 19. It shows that plants that have high cracking capacity enjoy higher profit margins. For every 10 percentage point increase in the cracking capacity of a refinery, the profit margin per barrel increases by 63 cents. This is because plants that have higher capacity are able to sell more products in gasoline market instead of the residual fuel market. I also find plants that have higher capacity in upgrade technology also enjoys higher profit margins per barrel, however it is not significant. The capacity in sulfur removal is found to have no effect in profit margin.

2.6 Conclusion

In this chapter, I generalize the existing CES production function to allow for plant characteristics and input differentiation. The generalized production function allows me to investigate how differences in plant technology or the type of inputs generates differences in the marginal cost and production possibility frontier. In the model, the type of crude oil is chosen endogeneously conditionaly the type of refinery and market prices. I estimate the multi-product production function using generalized method of moments. Using this approach I analyze how plant technology affects the profit margin of each plant. I find that plants that have higher capacity to transform crude oil such as cracking and upgrade enjoy higher profit margins per barrel.

2.7 Appendix

Table 13: Average Import percentage as total inputs from 1994 - 2009

East Coast	97.8%
Midwest	45.7%
Gulf	72.2%
Rocky Mountains	38.5%
West Coast	29.9%

Table 14: Estimation results on production function

product	API	Upgrade	Conversion	API*Upgrade	API*Conversion
gasoline	0.893*	-0.883*	-0.013	-1.446*	0.132
	(0.137)	(0.241)	(0.229)	(0.387)	(0.284)
jetfuel	2.829*	-0.682*	1.382	-0.784*	2.582*
	(1.532)	(0.272)	(3.571)	(0.259)	(0.732)
Diesel	3.483*	2.182	2.231*	0.973	-3.372*
	(0.712)	(1.948)	(0.842)	(0.583)	(1.264)
Residual Fuel	12.382*	5.485	11.943*	2.384	13.849*
	(3.584)	(9.172)	(2.382)	(1.934)	(2.489)
ρ	2.677*	(1.173)			

Standard errors are in brackets. * denotes significance under 5% confidence interval

Table 15: Estimated parameters on crude oil and natural gas ratio

	Natural Gas
API Difference	1.372(0.284)
Technology	0.983(0.534)
Sulfur Difference	6.382(1.948)
Technology	1.023(0.583)

Standard errors are in brackets. * denotes significance under 5% confidence interval

Table 16: Estimated parameters on Capacity Constraint

	Cost
Constant	2.018(0.843)
Utilization Rate	9.283(1.247)

Standard errors are in brackets. * denotes significance under 5% confidence interval

Table 17: Simulated output vs actual output ratio

product	simulated output	East Coast average
Propane	5.2%	3.7%
Gasoline	49.8%	51.8%
Jetfuel	8.2%	5.8%
Diesel	26.5%	30.5%
Fuel Oil	10.3%	8.2%

Using 2009 annual average price

Table 18: OLS Regression on Marginal Cost on Gasoline(excluding crude oil cost)

	Profit Margin
Constant	29.182(9.284)
Upgrade Technology	-11.530(4.236)
Cracking Technology	-18.274(3.038)
Desulfur Technology	1.202(0.847)

Standard errors are in brackets. * denotes significance under 5% confidence interval

Table 19: OLS Regression on Profit Margin

	Profit Margin
Constant	1.385(0.948)
Upgrade Technology	2.18(0.849)
Cracking Technology	7.382(1.849)
Desulfur Technology	-0.282(0.847)

Standard errors are in brackets. * denotes significance under 5% confidence interval

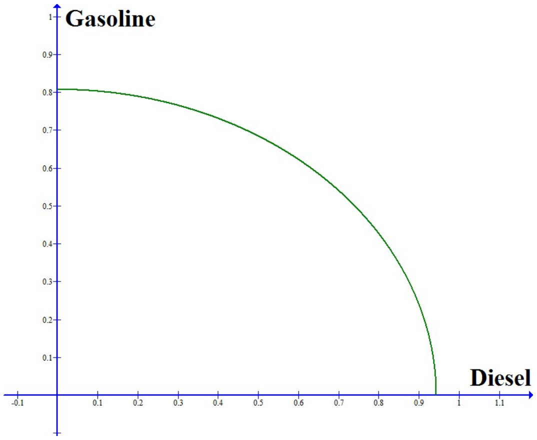

Figure 5: Production Possibility Frontier

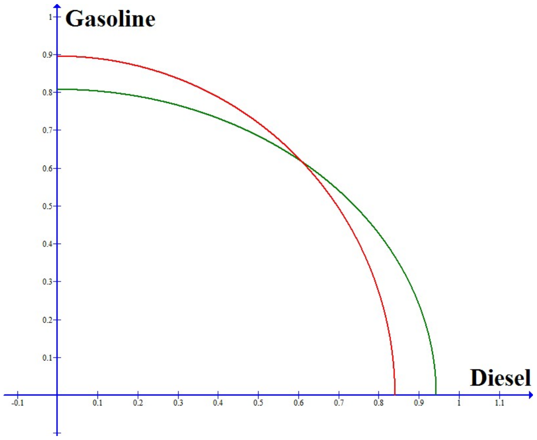

Figure 6: Production Possibility Frontier, Red curve is for lighter drude oil

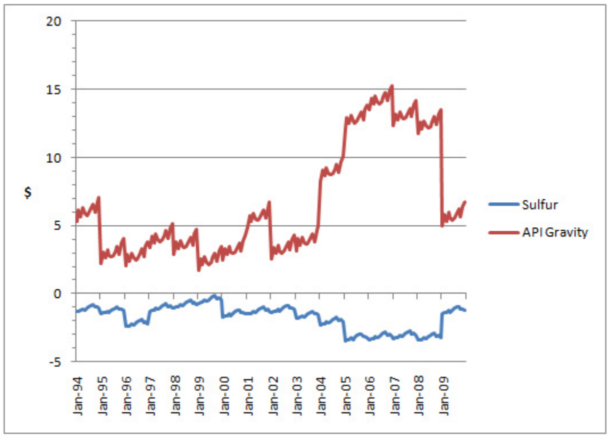

Figure 7: Change in Price per percentage change in API Gravity or Sulfur. Calculated at mean level

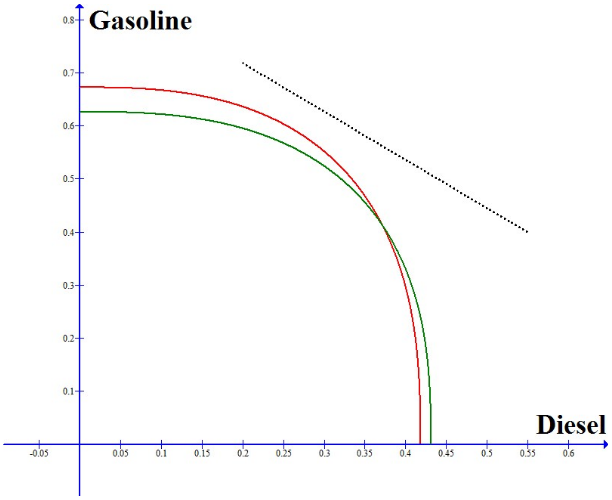

Figure 8: Green is for low conversion capacity. Red is for high conversion capacity. Dotted line is the average price ratio.

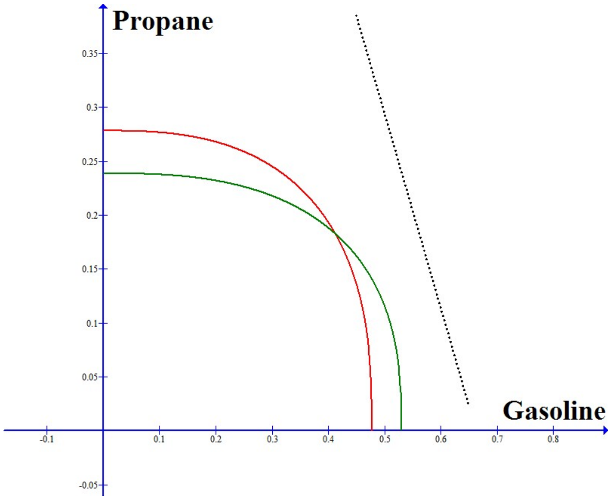

Figure 9: Green is for high upgrade capacity. Red is for low upgrade capacity. Dotted line is the average price ratio.

Chapter 3

The Effect of Sulfur Content Regulation on the U.S. Diesel Market

3.1 Introduction

In 2000, the U.S. Environmental Protection Agency decided to enact a policy change that reduced the amount of sulfur permitted in diesel fuel. The policy was implemented in 2006 and limited the maximum sulfur content in 80% of diesel outputs from large refinery to 15 parts per million. By 2010, the restriction was applied to all diesel outputs by U.S. refineries. Since all crude oil contains various amounts of sulfur, the policy change increased the burden of U.S. refineries to remove the sulfur from crude oil. The goal of this chapter is to quantify the effects of the policy change on the diesel fuel market. I investigate the impact on the marginal cost of diesel and other petroleum products for U.S. refineriers. Also, I quantify the effects on the diesel wholesale prices in each state and determine how each region in the U.S. may be impacted differently by the new regulation.

In order to investigate the implications of the policy change, it is important to consider the heterogeneity in refineries across the U.S. Unlike the traditional heterogeneity in total factor productivity, refineries produce a mix of outputs and each refinery has a comparative advantage for a different mix of products. Some plants are relatively more productive in producing lighter petroleum products such as gasoline, while other plants are more productive in producing heavier products such as diesel fuel. There are two reasons that explain the heterogeneity across plants. One is the difference in production technology across plants, the other is the difference in the type of crude oil available to each refinery. Both processing technology and the type of crude oil used for production dictate the mix of outputs that is achievable. Therefore availability to a particular type of crude oil is important to a refinery. Since each oilfield produces a distinct type of crude oil, the availability of various types of crude will depend on the location of an refinery in terms of its distance to the oilfields. For example, refineries on the East Coast will have different set of choices for crude oil available to them compared to a refinery located in the Rocky Mountains.

I use a CES production function to capture the substitutability between each output. The ratio of the output share depends on the technology of the plant and the type of crude oil used. Using parameters obtained from estimation, I run counterfactuals to investigate the implications of the policy change in each state. In order to run the counterfactual, I estimate state level demand functions for each of the outputs over the same time period. Then I use the estimated plant level production function and demand functions to calculate perfectly competitive equilibria under the old policy and under the full adoption of the new policy. Using results from the simulations, I quantity the effects of the new regulation by comparing the prices under the two simulated scenarios. The result shows that, on average across the U.S., the price increase is approximately 1.7%. However there is a large discrepancy in the increases between regions with high desulfur capacity and regions with low desulfur capacity. In regions with high desulfur capacity, the price increases are approximately half the increases in regions with low desulfur capacity.

There has been a large number of literature that uses reduced-form methods to estimate capture the treatment effect from a policy change. Much of these literature focused on the implications of environmental regulations on oxygenated gasoline, such as Brown et al. (2008)[4] and Borenstein and Shepard (2002)[3]. However, using the treatment approach in this scenario is problematic. Firstly, there is a lot of heterogeneity across plants in the technology of production processes. Secondly, the control group is not randomly selected, the exception for the new regulation was made for Rocky Mountain refineries and small refineries due to economic and cost reasons. Due to these two reasons, it is difficulty to investigate the policy effects on the untreated group.

3.2 Environmental Policy

In 2000, the EPA announced a comprehensive national control program to regulate heavy duty vehicle and its fuel as a single system. As a part of its program it aims to reduce the level sulfur in highway diesel fuel by 95%. There are two reasons for the reduction of sulfur in high way diesel fuel. Firstly sulfur emission from motor vehicles is a major contributor to acid rain, which causes both serious health and economical impacts across the U.S. The sulfur emission also results in fine air particles that causes breathing problems for many children and asthmatics. Secondly, the sulfur released also damages the newly required catalytic emission control devices in vehicles. These catalytic emission controls are used to reduce nitrogen oxides and other fine particles being released into the air which causes long term lung problems. The EPA hopes to reduce nitrogen oxides emissions by 90%. In order to achieve this result, it requires the simultaneous reduction of sulfur in diesel fuel.

Prior to the regulation change, the maximum sulfur level in diesel fuel was 500ppm, also known as Low Diesel Fuel (LDF). The program restricted refiners to start producing 80% of their diesel fuel with a sulfur content of no more than 15ppm, Ultra Low Diesel Fuel (ULDF), by June 1st of 2006. By June 2010, full conversion to ULDF was required for all refiners. Outline of the restrictions are in table 20. The program also provided hardship provisions to two types of refineries. It delayed the onset of the restriction to small refineries and Geographic Phase-in (GPA) Refineries. Small refiners are defined as refineries with less than 1500 employees and less than 155,000 barrels per day crude processing capacity and GPA refineries are refineries in the Rocky Mountains[1] and Alaska. Small refiners and GPA refiners were permitted to continue to produce LDF till 2010, by which they must also switch over to ULDF.

---

[1] States included are New mexico, Utah, Colorado, Wyoming, Idaho, Montana and North Dakota

3.3 Methdology

3.3.1 Demand

In order to calculate the counterfactual, I estimate the demand functions for each of the five petroleum products. Several sources of data are used for the estimation. For propane, the EIA provides monthly resale prices for each refining distrct. The EIA also provides monthly resale prices at state level for gasoline[2], jetfuel, diesel and residual fuel. The quantity data is also obtained from the EIA. The EIA releases wholesale level quantities at the state level for all five products. There are several states where no quantities are released. This occurs when there is only one refinery in the state and the quantity data is not released to protect plant level data being released. In this case, these states are lumped into other states as one market. For other explaining variables, the quarterly national GDP and yearly state gdp is obtained from Bureau of Economic Analysis. Vehicle registration data is obtained from the Highway Statistics Series published by the Office of Highway Policy Information. Data on aviation is obtained from the Bureau of Transportation Statistics.

I estimate a monthly level linear demand function for each state using instrumental variables. Two instruments are used and they both relate to spillage and accidents in petroleum product pipelines. One is the total value of losses due to an accident, the other is the total volume of petroleum products spilled due to an accident. Both data are available from the Pipeline and Hazardous Materials Safety Administratio. The results of the demand estimation and the explaining variables used are summarized in

---

[2] The EIA reports the average price across all gasoline grades, which is used in the demand estimation.

table 21. The five estimations are listed by column with the explaining variables on the left hand side. The remaining columns shows the estimated parameters using the appropriate explaining variables. In addition, I use yearly trend, monthly dummies and location dummies. The location dummy is interacted with the price variable in order to capture variations in price elasticity across states. All five estimations are regressed using 2 stage least squared with pipeline accidents as instruments for the price variable. The two instruments used are the quantity of fuel lost due to pipeline accidents and the value of damages caused by the pipeline accidents. In the bottom three rows, the adjusted R-squared, first stage F-stat and the price elasticity of demand for each of the five products are given. The instrument is very strong for gasoline, and slightly weak for jet fuel and diesel. However it is weak for propane and residual fuel. This is understandable because majority of the fuel transported through the pipelines are gasoline, jet fuel and diesel fuel. For the adjusted R-squared, it is very low for the estimation of propane. This is due to the lack of good explaining variables that captures the variation in demand for propane.

3.3.2 Refinery Objective

In the model, the maximum required sulfur level product j is denoted by $\bar{s}_j$ and the excess sulfur that needs to be removed for product j of refinery i is $s_i - \bar{s}_j$. The effect of the sulfur content requirement will impact the decision of the refinery through the cost function.

$$\begin{aligned} C(\mathbb{Q}_i) &= \min\nolimits_{z_i,a_i,s_i,NG} P(a_i,s_i)z_i + P^{ng}NG_i + \gamma_1(\tfrac{z_i}{k_{0i}})^{\gamma_2} \\ s.t.\ \ & \textstyle\sum_{j=1}^{J}(B_j(a_i,K_i)q_{ji}^{\rho})^{\frac{1}{\rho}}A_i = z_i \\ & NG_i = \textstyle\sum_j w(a_i,\bar{a}_j,s_i,\bar{s}_j,k_{3i})q_{ji} \end{aligned} \tag{3.1}$$

The sulfur policy impacts cost the of a refinery through the quantity of nautral gas that is required to remove the sulfur from the crude oil in the component $NG_i = \sum_j w(a_i,\bar{a}_j,s_i,\bar{s}_j,k_{3i})q_{ji}$. The function form for w_j which determines the quantity of natural gas needed per barral of petroleum product is

$$w(a_i,\bar{a}_j,s_i,\bar{s}_j,k_{3i}) = \epsilon_1(\bar{a}_j - a_i))_+^{\lambda_1}k_{i2}^{\lambda_2} + \epsilon_2(s_i - \bar{s}_j)_+^{\lambda_3}k_{i3}^{\lambda_4}. \tag{3.2}$$

The excess sulfur that is required to be removed is expressed through the term, $(s_i - \bar{s}_j)_+$. Therefore the amount of sulfur that needs to be removed vary depending on the product. The standard for gasoline is set nationally by the EPA and it is 300 parts per million prior to 2004, then the standard was changed to 30ppm by 2005. The standard for jet fuel is set by the FAA at 3000ppm. The standard for propane vary state by state, it is as low as 80ppm in California and as higher as 500ppm which is used in Wisconsin. In the simulation, the average of 300ppm is used, since propane makes up a very small percentage of ouputs and should not affect the results for gasoline and diesel fuel. There is also no specification for residual fuel since it is remnants of the crude oil after processing.

Given an output price for each market and each product p_{mj} and the price of inputs, I caulcuate the optimal output choices and input choices of the plant by solving the optimization problem of the plant by minimizaing the cost function 3.1 and maximizing the profit function

$$\begin{aligned} &\max_{\mathbb{Q}_i} \textstyle\sum_j p_j q_{ji} - \sum C(\mathbb{Q}_i) \\ &s.t. \quad q_{ji} \geq 0 \ \ \forall j \end{aligned} . \tag{3.3}$$

In order to obtain the total factor productivity for refineries that were not part of the initial estimation, I use the production function parameters and cost function parameters obtained from the intial estimation and calculate the residuals for the remaining refineries. I observe the aggregate input used in a refining district, $\sum_i z_i$. I observe the parameters and type of crude oil used by each refinery. This allows me to calculate the optimal output quantity of each plant by solving the profit maximizing problem described in equation 3.3. Then I calculate the mean total factor prodocutivity shock using

$$A \sum_i \sum_{j=1}^{J} (B_j(a_i, K_i) q_{ji}^{\rho})^{\frac{1}{\rho}} = \sum_i z_i \tag{3.4}$$

3.3.3 Market Equilibrium

After determining output of each refinery, I calculate the quantity sold in each market. The distribution of the output to each market is calculated using observed data. Using data on the flow of petroleum products between PADDs in the US that is provided by the EIA, I calculate the share of outputs a refinery in state m produces that flows into each state m', denoted by $\frac{q_{m'j}}{\sum_{m'} q_{mm'j}}$. Using the distance between the refinery and the market, and the percentage of a plant's output that flows into that market, I estimate the discount rate conditional on the distance.

$$\frac{q_{m'j}}{\sum_{m'} q_{mm'j}} = \frac{exp(-\alpha dist_{m'})}{\sum_{m'} exp(-\alpha dist_{mm'})} \tag{3.5}$$

. After obtaining the discount rate α, I use it to calculate the quantity sold in each market. This method assumes that refineries have little flexibility in adjusting state level shipments using real-time feedback. This assumption is justified since refineries need prebook the transportation of petroleum products to wholesalers based on pipeline schedule.

After obtaining the quantity sold in each market, the new market price is calculated based on the petroleum products demand. The new market price is then used to recalculate the output of each refinery. This step is repeated until the market clearing price for each state is reached. In order to simulate the counterfactual, I still require the demand for each product in each market and the total factor productivity, the term A_i in equation 3.3, for regions that were not used for the estimation in chapter 2.

I run the simulation twice under difference scenarios. The first simulation is conducted under the pre-2006 policy using June 2006 prices and plant characteristics. The second simulation is conducted using the same prices and plant characteristics but with a full implementation of the policy with the 15ppm restriction applicable to all refineries. Then the simulated price and cost are compared between the two scnearios.

3.4 Result

Table 22 compares the change the marginal cost under difference scenarios. Column 2 shows the difference when refineries are not allowed to alter the type of crude oil. Column 3 shows the charge in marginal cost when refineries are allowed to alter the type of crude oil but the distribution of outputs is fixed. In column 2, without considering how firms may endogeneously alter their behavior, it shows a 9.2 dollar per barrel increase for diesel fuel and no change for other outputs. This overestimates the impact of the policy on the cost of diesel fuel by not allowing producers to compensate by altering the distribution of outputs and the type of input. Comparing between column 2 and column 3, it shows that, with the policy implementation, diesel fuel stil has the greatest price increase of 4.2 dollars per barrel. However it is lower than the previous scenario by 5 dollars per barrel. The marginal cost of other outputs increases as well. There are two forces here that affect the marginal cost of other outputs. Due to the policy change, refineries decrease the sulfur level in crude oil. This decreases the processing cost of other outputs, but it increases the cost of the crude oil itself. Column 4 shows the result when both the type of crude oil and the output quantities becomes endogeneous. Due to the policy change, refineries decreases their production in diesel and shift to other products such as gasoline. As a result, it further reduces the impact of the policy change on the marginal cost of diesel production. On average, the policy causes an increase of 2.6 dollars per barrel in the marginal cost of diesel production.

Furthermore I examine how does the change in marginal cost vary depending on the technology of the plant. I regress the change in marginal cost of each plant on the capacity of the plant processing technologies and the result is show in table 23. It shows that plants that have higher desulfur technology has lower increases in marginal cost for diesel fuel than plants with lower desulfur technology. A 10% increase in capapcity to remove sulfur is coorelated with a reduction of 35 cents per barrel in the increase of marginal cost for diesel fuel. Also plants that have higher capacity for conversion has lower increases in marginal marginal cost. The same result also holds for the marginal cost of gasoline.

I also examine the effect of the policy on the reduction of sulfur used in production. I regress the simulated reduction in sulfur content on regional dummies and plant technology. The result of the regression is presented in 24. It shows that plants along the East coast and and Gulf had the greatest decrease in the amount of sulfur in crude oil. Controlling for everything else, plants in East Coast reduced the amount of sulfur in crude oil by 0.71 percentage points. The result also shows that refineries that have higher capacity for desulfur reduced sulfur by a lower amount compared to refineries that have lower capacity for desulfur.

Finally, the change in the simulated price in the diesel fuel market is reported in figure 10. The color of the states indicates the level of price increase. Darker blue states have higher increase in price and lighter blue states have a lower increase in prices. Yellow dots indicate a refinery with low desulfur capacity and red dots indicate a refinery with higher desulfur technology. The increase in price due to the full implementation of the policy ranges from 2.6 cents per gallon to 5.1 cents per gallon. This implies a percentage increase in price from 1.21% to 2.34%. The map shows that in areas with higher concentration of high desulfur capacity plants such as Texas, Midwest and Wyoming, the price increase is lower compared to regions with low desulfur capacity plants. The figure shows greater price increases in the New England area and in Florida. The main reason for the higher price increases in these two regions is the lack of refineries in those areas. These two regions rely on petroleum products processed in other regions.

3.5 Conclusion

In this chapter, I use a multi-product production function model to quantify the impact of the EPA's lowering of maximum sulfur level in diesel fuel on petroleum products market. I allow the choice in the type of the crude oil to depend on the technology of each plant and in the availability of crude oil that varies with the location of the plant. This is important because the policy change will impact the type of crude oil used by a refinery and the response of each plant will differ according to each plant's technology and location. Due to the heterogeneity across plants, in order to investigate the impact of the policy, I use a model that takes into account the differences in technology and allows the plant to choose the type of crude oil used in response to the policy change.

The result shows plants that have a lower level of technology will be more negatively impacted by the policy change. The increase in marginal cost for plants at the bottom 20 percent quantile in desulfur technology is twice as much as plants at the top 20 percent quantile in desulfur technology. I also conducted simulations to quantify the impact of the policy change on prices. If the same policy was implemented across all U.S. refineries instead of just large refineries outside of the Rocky Mountains in 2006, it would have resulted in a 3.4 cents per gallon increase in the price of diesel fuel on average across the U.S. Furthermore, in the Rocky Mountains area, the price increase would have been approximately 20% higher compared to other regions of the U.S. This is due to the lower levels of sulfur removal technology for plants in the Rocky Mountains during 2006 and more costly access to low sulfur crude oil. This result raises policy implications. Heterogeneity in technology and crude oil access across plants plays an important role in the response of the diesel market to the policy change. In regions with a lower technology or more costly access to low-sulfur crude oil, the impact of the policy is much greater compared to regions with a higher level of technology or cheaper access to low-sulfur crude oil. As a matter of fact, when the policy was decided in 2000, the lowering of sulfur limit for Rocky Mountains refineries were delayed by 4 years from 2006 to 2010. The result of the chapter suggest that if the exception was not made, consumers in the Rocky Mountains woul have faced almost a 2% increase in the price of diesel compared to an approximately 1.5% increase in price for the rest of the country.

3.6 Appendix

Table 20: Sulfur Content Regulation Change Summary

Date	Large Refiners	Small Refiners and GPA Refiners
< 2006, June	500	500
2006, June - 2010, June	80% 15, 20% 500	500
> 2010, June	15	15

Unit is parts per million.

Table 21: Demand Estimation Average

	Propane	Gasoline	Jet Fuel	Diesel Fuel	Residual Fuel
Price	-1.38	-3.49	-2.19	-2.43	-1.32
Natl GDP	0.00	0.004	0.001	0.002	0.002
Temp	-0.01			-0.03	
Vehicles		0.002			
State GDP	0.000	0.003	0.001	0.000	0.001
Airline Passengers			0.98		
Airline Miles		1.94			
Heavy Vehicles				0.02	
Adj R-Squared	0.24	0.69	0.73	0.72	0.41
1 stage F-stat	4.38	11.43	7.61	8.51	5.82
Price Elasticity	-0.76	-0.54	-0.62	-0.47	-0.32

All estimations include year trend, monthly dummy and location dummy. The price elasticity is also interacted with location dummy.

Table 22: Simulated change in marginal cost across refineries

Product	Observation	Fixed Input	Endogeneous Input Type	Full Simulation
Propane	142	0	0.2	1.4
Gasoline	142	0	0.6	1.5
Jetfuel	142	0	0.9	1.3
Diesel	142	9.2	4.2	2.6
Residual Fuel	142	0	2.7	1.6

Unit is dollars per barrel. Calculated under 2006 June input prices

Table 23: OLS regression of simulated marginal cost change on plant technology

Technology	Gasoline	Diesel
Upgrade	-0.08	-0.2
	(0.07)	(0.04)
Conversion	-1.2	-2.2
	(0.4)	(0.9)
Desulfur	-0.7	-3.5
	(0.3)	(0.4)
Adj R-Squared	0.38	0.41

Table 24: OLS regression of simulated sulfur content reduction in crude oil on plant technology

Variable	Sulfur in crude oil
PADD 1	-0.71
	(0.16)
PADD 2	-0.13
	(0.22)
PADD 3	-0.68
	(0.21)
PADD 4	-0.11
	(0.14)
PADD 5	-0.31
	(0.16)
Upgrade	-0.04
	(0.11)
Conversion	0.09
	(0.11)
Desulfur	0.17
	(0.08)
Adj R-Squared	0.56

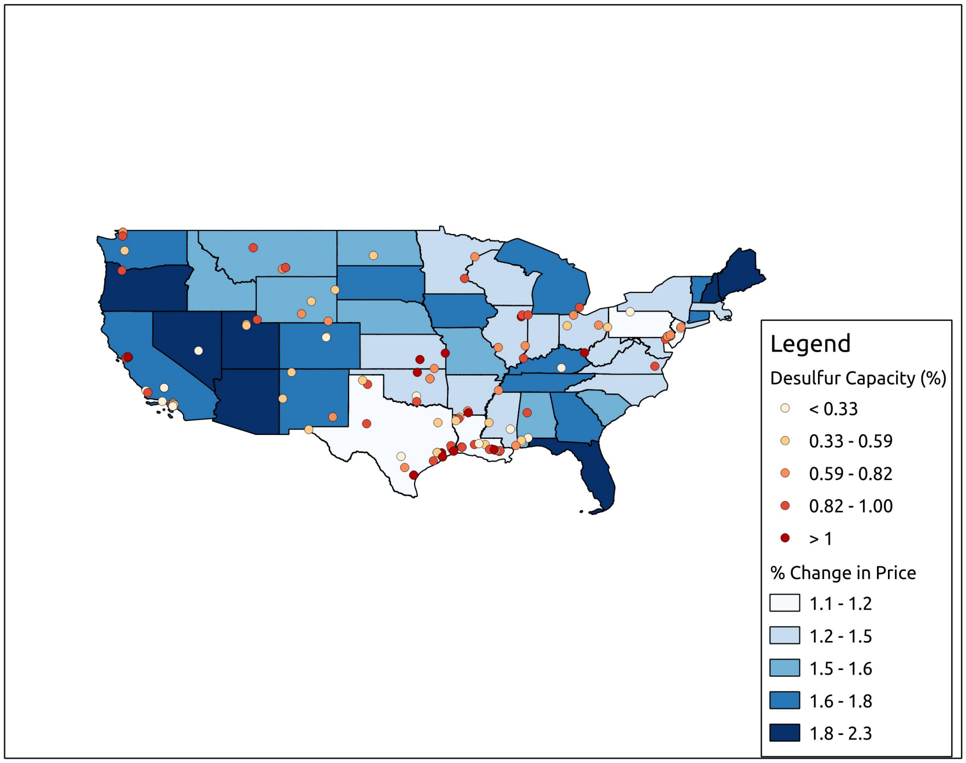

Figure 10: Change in Diesel Prices

Bibliography

[1] K.J. Arrow, H.B. Chenery, B.S. Minhas, and R.M. Solow. Capital-labour substitution and economic efficiency. *Review of Economics and Statistics*, pages 225–250, 1961.

[2] Auffhammer, Maximilian, and Ryan Kellog. Clearing the air? the effects of gasoline content regulation on air quality. *American Economic Review*, 101(101):2687–2722, October 2011.

[3] Severin Borenstein and Andrew Shepard. Sticky prices, inventories, and market power in wholesale gasoline markets. *Rand Journal of Economics*, 33(1):116–139, 2002.

[4] Jennifer Brown, Justine Hastings, Erin T. Mansur, and Sofia B. Villas-Boas. Reformulating competition? gasoline content regulation and wholesale gasoline prices. *Journale of Environmental Economics and Management*, 55:1–19, 2008.

[5] Curtis Carlson, Dallas Burtraw, Maureen Cropper, and Karen L. Palmer. Sulfur dioxide control by electric utilities: What are the gains from trade? *Journal of Political Economy*, 108(6):1292–1326, December 2000.

[6] Matthew Chesnes. Capacity and utilization choice in the us oil refining industry. 2009.

[7] Hayley Chouinard and Jeffrey Perloff. Gasoline price differences: taxes, pollution regulations, mergers, market power, and market conditions. *The B.E. Journal of Economic analysis & Policy*, 7(1):8, 2007.

[8] Meredith L. Fowlie. Emissions trading, electricity industry restructuring, and investment in pollution control. *American Economic Review*, 100, 2010.

[9] Giora Hanoch. Polar functions with constant two factors - one price elasticities. In Melvyn Fuss and Daniel McFadden, editors, *Production Economics: A Dual Approach to Theory and Applications*. North-Holland, 1978.

[10] Sudhanshu Kumar Mishra. A brief history of production functions. *Social Science Research Network Working Paper Series*, October 2007.

[11] Erich J. Muehlegger. Gasoline price spikes and regional content regulation: a structural approach. 2006.

[12] Y. Mundlak and A. Razin. On multistage multiproduct production functions. *American Journal of Agricultural Economics*, 53(3):491–499, 1971.

[13] Yair Mundlak. Specification and estimation of multiproduct production functions. *Journal of Farm Economics*, 45(2):433–443, May 1963.

[14] Yair Mundlak. Transcendental multiproduct production functions. *International Economic Review*, 5(3):273–284, September 1964.

[15] C. Richard Shumway. Estimation of multicrop production functions: Comment. *American Journal of Agricultural Economics*, 70(3):729–732, August 1988.

[16] H. D. Vinod. Econometrics of joint production. *Econometrica*, 36(2):322–336, April 1968.

[17] R.D. Weaver. Multiple input, multiple output production choices and technology in the u.s. wheat region. *American Journal of Agricultural Economics*, 65(1):45–56, 1983.