How Shocks to Input and Output Markets Change Spreads and Asset Value: The Case of the Refinery Industry

Preliminary and incomplete. Please dont quote

Hamed Ghoddusi, Sheridan Titman, Stathis Tompaidis *

June 13, 2011

*Ghoddusi is with the Vienna Graduate School of Finance, Hamed.Ghoddusi@vgsf.ac.at. Titman is with the McCombs School of Business, University of Texas at Austin, Department of Finance, Sheridan.Titman@mccombs.utexas.edu. Tompaidis is with the McCombs School of Business, University of Texas at Austin, Department of Information, Risk and Operations Management and Department of Finance, Stathis.Tompaidis@mccombs.utexas.edu.

How Shocks to Input and Output Markets Change Spreads and Asset Value; The Case of the Refinery Industry

ABSTRACT

We present an equilibrium model of price dynamics and the transmission of shocks in a supply chain. Starting with exogenous processes for the net supply of the upstream input and the demand for the downstream output, we construct the equilibrium process for the input and output prices, the spread between input and output prices, and the value of the capital asset that transforms the input into the output. We specify and calibrate our model for the case of the crude oil and gasoline in the context of oil refineries and estimate the structural parameters. Moreover, we provide comparative statistics from our model and empirical evidence supporting the predictions.

JEL Classification: G13, G32, Q43

Introduction

In its simplest form, a supply chain consists of three components: an upstream primary input, a downstream output product, and a capital asset that turns the input into the output. One can give numerous examples: crude oil is transformed into refined products at a refinery; electricity is generated through converting different fuels; bauxite ore is transformed into aluminum in aluminum smelters; and raw tomatoes are converted into paste. In each case, raw materials are processed into a final commodity.

In this paper, we present an extended partial equilibrium model that characterizes processes for input and output prices, as well as the spread between these two prices. The model accounts for the fact that the supply of the input good as well as the demand for the output good are subject to stochastic shocks, and both type of shocks can lead to changes in input and output prices. As we illustrate, the capacity constraint of the converting asset and the cost structure of the production function determine the extent to which input supply and output demand shocks translate into price changes in the other market. We characterize the production function, and show how the convexity of the total cost function, as well a positive cost for deviation from the historical production level, influence both price processes and their spread.

Our primary interest is on the level and volatility of spreads between the input and output prices, which determine the profitability of the producer. As our model illustrates, capacity constraints lead into two major production regions: an interior region in which prices of input and output are highly connected to each other, and spreads are small; and a boundary region in which the supply is inelastic, and input and output prices are decoupled. In addition to capacity constraints, market structure also plays a role. The spread between the input and output prices decreases with the elasticity of the demand for the output; the spread, also, increases with the convexity of the production process. Additionally, the spread tends to be more volatile when the producer is a monopolist.

We apply our model to the US oil refinery industry, where the input is crude oil and the outputs include refined products like gasoline, jet fuel, and heating oil. The model generates several testable implications: a) in a perfectly competitive market with no constraints in the operation of the capital asset spreads converge to a fixed level. However, if the operation of the capital asset is constrained — through either capacity limits, or production adjustment costs — spreads deviate from this level. In a static set–up, spreads are always positive, while a dynamic behavior (emerging from adjustment costs or storage) allows for negative spreads in some periods; b) The distribution of cash flows accruing to the capital asset is highly skewed to the right; it is near zero or even negative for a significant fraction of time, but jumps to a large positive level in certain periods. Attempting to eliminate these rare moments of high profitability (e.g. through regulation or taxation) lowers the incentive to invest in capital assets; c) the volatility of the spread between the input and output prices is positively related to the level of demand of the output: high demand pushes production to the inelastic production region in which input and output prices are decoupled, leading to larger and more volatile spreads; d) when demand for the output is high enough, spreads between input and output prices under both monopolistic and competitive market structures converge; e) very low and decreasing input prices, and stable output demand increase spreads. In this case, the capital asset operates at nearly full capacity; f) increasing the volatility of input supply and output demand increases the average spreads by increasing the frequency at which the capital asset produces at full capacity; this effect declines when the correlation between shocks to input and output markets increase; g) The correlation between the input and output prices depends on the levels of output demand and input supply. When output demand and input supply are low, the prices of the input and the output are highly correlated. When output demand is strong or input price is very low, the prices of the input and the output diverge, and the correlation drops; h) The correlation between input and output prices is higher in competitive markets than in monopolistic markets.

We perform empirical tests for some of our theoretical results, by looking into the effect of input/output price levels and volatilities on the value of the refinery industry. We run a contemporaneous return regression by regressing the stock return of refining firms on a group of explanatory variables. Controlling for the usual return factors of Fama–French, we show that the ratio of gasoline and crude oil prices in spot and forward markets have a significant and positive effect on the stock price of a refinery. Furthermore, the volatility of input price and the excess return of crude oil have positive and significant effects. Robustness checks, provide support to the hypothesis that these are refinery-specific cash-flow factors, not proxies for some omitted return factors.

Finally, we calibrate our dynamic model using the simulated method of moments (SMM) approach, by fitting the moments of the model-generated price and quantity time-series to their empirical counterparts. We conduct this exercise in both spot and forward markets to estimate parameters under the physical as well as the risk–neutral measures. The resulting parameter values, paired with a structural dynamic model, enable us to perform potential policy analysis as well as contingent claim pricing.

We are, of course, not the first to consider the interaction of capacity constraints, market volatility, and the asset value. Aristotle described what might have been the first discussion on the effect of input supply on the profit of an industry, which has fixed capacity in the short–term. In the book of Politics, Aristotle describes how Thales of Miletus correctly forecasted a good year for olive production in ancient Greece, and based on this forecast, secured the use of all the local olive presses. The subsequent realization of a good harvest resulted in an increased demand for the limited number of olive presses, and allowed Thales to rent out the olive presses for a large profit.

Within the context of our model, we see that an increase in the supply of olives will result in olive presses operating at full capacity, an increase in the spread between the price of olives and the price of olive oil, and a subsequent increase in the equilibrium rental rate on olive presses. Aristotle attributes Thales' large profit to his monopoly position, which would of course contribute to such profits. But as our model illustrates, the incremental increase in profit due to the good olive harvest would actually be the same under perfect competition for a fixed number of olive presses, as long as the olive presses operate at capacity.

More recent papers on this issue include **?**, **?**, and **?**, among others, who build structural models of the electricity sector. **?** looks into the refinery industry but has a different focus on investment and maintenance issues. In addition to structural models, a growing line of literature in option pricing focuses on pricing of spread options with crucial and realistic features of co–integration and GARCH effect (see **?** for a good review, also **?**, and **?**). Our paper differs from the current literature in several important ways. First, we build a dynamic model, which is more detailed in terms of technological specifications, and we calibrate it using spot as well as futures prices. Second, we provide empirical evidence supporting our theorem. Third, we have a special focus on the effect of volatility and adjustment costs on asset value and their interactions with capacity constraints.

The remainder of our paper is structured as follows. Section **??** provides basic information about the refinery industry. In Section **??** we present the structural model for the upstream input and downstream output markets of the supply chain as well as assumptions on the operation of the capital asset. Sections **??** and **??** discuss the behavior of the model, and its consequential propositions. Section **??** provides the results of several econometrics tests on the theoretical predictions of our model. Section **??** calibrates our model for the case of refineries with crude oil as input, and gasoline as output. Section **??** discusses potential applications of our model and concludes the paper.

I. The Refinery Industry

The refining industry produces an array of products, including gasoline, distillates – i.e. diesel fuel, jet fuel, and heating oil – and heavy or residual products, such as gas oil, lubricants, and asphalt. However, gasoline by far is the most important product for US refineries. Gasoline yield (the ratio of gasoline output to total refinery output) has been moving between 42% and

48%, and due to higher gasoline prices, its share in refinery revenue is around 60%. Appendix **??** provides information regarding the average yield of various refinery products for the period of 1993–2010.

As an industry that converts one type of commodity into other types, the refinery industry benefits from price differentials in the output and input markets. The standard term of "crack spread" refers to the difference between the price of certain derivatives of crude oil (mostly gasoline and heating oil) and the crude oil used to produce these refined products. The refining margin, which is the major driver of value in refinery industry, can be proxied by crack spread[1]. There are synthetic contracts traded in NYMEX and other commodity exchanges that offer a 3-2-1 ratio, meaning that the value of the contract is the value of the sum of two units of gasoline and one unit of heating oil, net three units of crude oil.[2] One of the goals of our paper is to understand the dynamics of the value of the crack spread in a structural equilibrium model.

The descriptive statistics of crack spreads for the interval of 1986–2010 are provided in Table **??**. The values for the Skew and Kurtosis of this process suggest that the distribution is not symmetric, and that there is a higher concentration of large values on the right side of the distribution. Our model explains this feature very well.

A key empirical puzzle that has received a significant amount of attention is asymmetric response of gasoline retail prices to crude oil shocks. The asymmetry refers to the fact that a positive shock to crude oil market transmits much quicker to gasoline prices than negative shocks do. **?** explain why gasoline prices do not immediately react to the changes of crude oil price. Their model includes production adjustment costs and the possibility of storage. They use the response of gasoline futures prices to the innovations of crude oil futures with the

[1]Some US refiners that can process cheaper heavy or sour crude oil benefit significantly from the price differentials between these types and light/sweat crude oil. When the oil market becomes tight or producers of heavy/sour crude change their production policy, the price gap between types shrinks and they lose this opportunity. We do not model this heterogeneity in crude oil prices

[2]Although such a contract can easily be replicated by making a portfolio of individual contracts on these three underlying products, market participants prefer the standard crack spread contract because of its lower margin account requirements.

Statistics	Value
Mean	17.9
Median	13.9
Maximum	107.2
Minimum	2.4
Std. Dev.	11.9
Skewness	2.3
Kurtosis	11.2

Table 1
Descriptive Statistics of Gasoline Crude Oil Crack Spreads

same maturity. If the lag between gasoline and crude oil prices is due to the slow adjustment of production, then near-to-maturity contracts of gasoline should react only partially, whereas long maturity contacts will react immediately with a full adjustment. Their results show that a 1–cent increase in crude oil price eventually increases gasoline prices by 1.14 cents, but this effect is 0.16 cents less for near-month futures contracts. Also, **?** examines whether the volatility of crude oil price affects the degree of asymmetric response of gasoline prices to oil price shocks. His findings are based on a VAR model and confirm that there is a negative relationship between volatility and the degree of price response asymmetry. He suggests that these findings may support a search model with Bayesian updating and oligopoly coordination in the gasoline market.

Major players pf gasoline market have different industrial organizations. Few companies are involved only in the refining business, and purchase their crude oil from the market. Whereas, some vertically integrated ones control the entire chain from the upstream to retail markets. **?** show that oil companies that own refineries may use strategic pricing of crude oil at regional level to influence the input cost of their rivals.

The degree of competition in the refinery industry is a key issue for our modeling purpose. **?** estimate Herfindahl-Hirschman Index (HHI) for five "Petroleum Administration for Defense Districts" (PADDs) and conclude that, except for the East Coast, HHI is in the boundary of competitive markets in all other four one. Even on the East Coast, it can not be concluded that gasoline market is not competitive because this region imports a lot of gasoline from other regions. **?** also reports similar results for New York, Gulf Coast, and Los Angeles. Following these results, we will model the refinery industry as a competitive case.

II. Model

We develop an equilibrium model for the price of the input, the output, and the value of the capital asset that transforms the input into the output. The model is built on assumptions on the supply of input, the demand of the output, and the operational characteristics of the capital asset. We express our model for the case of crude oil as input, gasoline as a representative output of a refining process, and a refinery as the capital asset that transforms the input into the output. However, our model is general and can be applied to other examples of supply chains.

A. Input: Crude Oil

We assume that the supply of the input is in competitive equilibrium, implying that the price of the input equals the marginal cost of its production. To account for the increase in the marginal cost of oil extraction, given increases in the aggregate level of production, we assume that the marginal cost is linear in aggregate quantity

$$P_C = \Delta + \alpha Q_G \tag{1}$$

where P_C is the price of crude oil, Δ the net supply of crude oil not used to produce gasoline, Q_G the quantity of crude oil used to produce gasoline, and α the constant elasticity of marginal cost of crude oil production with respect to the aggregate quantity produced.

Global crude oil supply is determined by the production of OPEC countries, large producers such as the United States and Russia, and a competitive fringe of small producers. Two major explanations support the assumption of mean–reverting behavior, especially in the medium–term time horizon. First, in a resource production economy with optimal investment (e.g. **?**, **?**, and **?**), prices will be mean–reverting. Second, OPEC's trigger–type policies induce mean–reverting price behavior. Their policies imply that supply increases when prices cross a certain cap level and decreases when prices drop below a floor level, leading to crude oil prices that exhibit mean-reversion. We capture this effect in our model by modeling the net non-US demand effect on crude oil price as a mean-reverting process.

$$d\Delta = \mu_\Delta(\overline{\Delta} - \Delta)dt + \sigma_\Delta dW_C \tag{2}$$

Here the mean reversion rate μ_Δ, the long term level $\overline{\Delta}$, and the volatility σ_C are assumed constant. The process Δ represents the tightness in the supply capacity of global crude oil market. When non-US demand jumps up, Δ increases and pushes the equilibrium crude oil price to a higher level.

B. Output: Gasoline

The price of gasoline is determined by the demand factor X^f, and the aggregate supply of gasoline to the market Q_G. The stochastic and seasonal demand factor X^f depends on long–term variables like the fuel–efficiency of the stock of vehicles in use and short–term shocks to income, weather, and travel seasons.

Since gasoline has a lower density, one volume unit of crude oil produces $c_{C,G}$ ($c_{C,G} > 1$) volume units of gasoline. However, for simplicity of notations throughout the paper, our unit of gasoline is the amount produced from one barrel of crude oil.

The dynamics for the deseasonalized demand factor X follows a mean-reverting process with shocks that are correlated with the shocks in the net supply of crude oil. The correlation comes from the fact that business cycle shocks of the US economy and world economy are correlated.

$$\begin{aligned} X^f &= X f_g \\ X &= \mu_G(\overline{X} - X)dt + \sigma_G dW_G \end{aligned} \tag{3}$$

Where f_g is the seasonality factor, μ_G the speed of mean reversion, $\overline{X}$ the long term level, and σ_G the volatility of the shocks to the deseasonalized gasoline demand. The correlation between the shocks of gasoline demand and crude oil supply, ρ, is assumed constant

$$dW_G dW_C = \rho dt \tag{4}$$

The demand curve for gasoline is assumed to be a linear function

$$P_G = X - b_G Q_G \tag{5}$$

Where b_G is a time invariant elasticity parameter. While a linear demand function allows the existence of a price above which demand for gasoline is zero, we will use it as an approximation of the true demand function. However, in the calibration section we will be more flexible and will adopt different functional forms.

C. Capital Asset: US Refinery Industry

We assume that all the gasoline consumed in the US is produced domestically, and that there are no gasoline imports or exports.[3]

The domestic market is supplied by a competitive refinery industry that converts crude oil into refined products including gasoline, heating oil and jet fuel. In reality, refineries may use a few different forms of crude oil — e.g. heavy crude and light sweet — depending on their engineering design and technical specification. For simplicity, we do not distinguish between different input types, and we treat crude oil as a homogeneous commodity. Although refineries can make adjustments in their refined products mix, historical yield data show that the share of gasoline, heating oil and jet–fuel from the total productions is approximately 80%. We abstract from the product-portfolio optimization problem of the refinery and simply assume an aggregate demand and product model.

C.1. The Cost of Gasoline Production

Production of gasoline involves explicit factor costs and implicit organizational and depreciation costs. The total cost function of gasoline production, $TC(Q)$, for a representative refinery is given by

$$\begin{aligned} \mathrm{TC}(Q_G,\Delta,Z) &= F_C + P_C(Q_G,\Delta)Q_G + P_{I,G}Q_G + \frac{\phi Q_G^2}{2} + \frac{\delta}{2}(Q_G - Z)^2 \\ Z_t &= \eta \int_0^t e^{\eta(u-t)} Q_{G,u} du \end{aligned} \tag{6}$$

where F_C is the sunk fixed cost, and Z_t is a weighted average of the historical gasoline production.

[3] The actual imports of gasoline to the US market is not large compared to the total production. For states like California that have rather stringent environmental standards for gasoline, the import is practically zero. The import is more relevant for the East Coast where some gasoline is imported from the UK and Europe. We will account for the import of gasoline when calibrating our model.

The variable part of the total cost function consists of four major cost elements:

- the unit cost of crude oil which depends on gasoline and non-gasoline demand and supply shocks, $P_C(Q_G, \Delta)$,
- the cost of other inputs except crude oil which includes items such as energy (especially natural gas), labor, maintenance and chemicals used in the process of turning crude oil into gasoline[4], $P_{I,G}Q_G$
- the cost associated with producing at high levels of capacity utilization. This cost may occur because the refinery may experience higher probability of break-down in later periods (see **?**), may require over-working of its labor force, higher payments to contracts, and because the marginal units used to produce at high levels might be more costly to build. We model the capacity-related costs using a quadratic form $\frac{\phi}{2}Q_G^2$ where ϕ is a constant coefficient.
- Finally, adjusting the production rate to a level different from the average historical level is costly because it requires sudden changes of production plans and re-allocation of resources. **?** estimates the value of adjustment costs of changing production portfolio for refiners and finds a non–zero adjustment cost. Following **?** we use a quadratic adjustment cost, $\delta(Q_G - Z)^2/2$ with constant coefficient δ.

D. Optimization Problem

The production of gasoline requires a quasi–fixed capital input and a set of flexible inputs. The investment in capital determines the maximum production rate $\overline{Q}_G$ which is given exogenously. We abstract from the random shocks to capacity, resulting from natural disasters and industrial accidents, and assume a fixed capacity.

[4]The price of emission allowances is becoming relevant for refiners in countries that implement emission reduction policies. In future, one might need to add the allowance price as the second major stochastic cost element

The representative firm of the competitive industry solves a dynamic optimization problem to maximize the total social surplus (TSS) [5] in input and output markets. The value function depends on three state variables: the historical average production Z, the level of gasoline demand X, and the net-of-gasoline demand supply in the global crude oil market Δ. We assume that the time-invariant "market price of risk" for the refinery industry is exogenously given. Therefore, the industry uses the rate r to discount the future streams of profits.

$$
\begin{aligned}
&V(Z,X,\Delta) = \mathrm{Max}_{Q_G \leq \overline{Q}_G} E \int_0^\infty \left[\mathrm{TSS}(Q_{G,t}, X_t, \Delta) - P_{I,G} Q_{G,t} - \frac{\phi Q_{G,t}^2}{2} - \frac{\delta}{2}(Q_{G,t} - Z_t)^2 \right] e^{-rt} dt \\
&\mathrm{TSS}(Q_G, X_t, \Delta) = \int_0^{Q_G} (P_G(Q_{G,u}, X_t) - P_C(Q_{G,u}, \Delta)) du \\
&dZ_t = \eta(Q_{G,t} - Z_t) dt \\
&dX_{G,t} = \mu_G(\overline{X} - X_{G,t}) + \sigma_G dW_G \\
&d\Delta = \mu_\Delta(\overline{\Delta} - \Delta) dt + \sigma_\Delta dW_C
\end{aligned}
\tag{7}
$$

The value function, V, satisfies the following Hamilton-Jacobi-Bellman (HJB) equation

$$
\begin{aligned}
&rV = \mathrm{Max}_{Q_G \leq \overline{Q}_G} \Big[\mathrm{TSS}(Q_G) - P_{I,G} Q_G - \frac{\phi Q_G^2}{2} - \frac{\delta}{2}(Q_G - Z)^2 + \eta(Q_G - Z)V_Z + \mu_G(\overline{X} - X)V_X + \\
&\frac{1}{2}\sigma_G^2 V_{XX} + \mu_\Delta(\overline{\Delta} - \Delta)V_\Delta + \frac{1}{2}\sigma_\Delta^2 V_{\Delta\Delta} + \rho \sigma_\Delta \sigma_X V_{X\Delta} \Big]
\end{aligned}
\tag{8}
$$

Considering the fact that the optimal policy is restricted, the first order condition leads to

$$
X_G - b_G Q_G - \Delta - \alpha Q_G - P_{I,G} - \delta(Q_G - Z) - \phi Q_G + \eta V_Z + \lambda(\overline{Q}_G - Q_G) = 0 \tag{9}
$$

Here λ is the Lagrange multiplier associated with capacity constraints. In addition to the excess cost of producing at higher capacity utilization, the producer takes into account two dynamic cost elements: the instantaneous cost of deviating from the historical production level, $\delta(Q_G - Z)$, and the future cost of choosing a production level that changes the historical

[5] Total social surplus is the area between the output demand and input supply curves. By the first welfare theorem, solving the problem of a competitive firm is equivalent to maximizing TSS

average, ηV_Z. When the optimal production of gasoline is less than the maximum capacity, the Lagrange multiplier associated with the capacity constraint is equal to zero. In that case, the optimal quantity is given by

$$Q_G^E = \frac{X - P_{I,G} - \Delta + \delta Z + \eta V_Z}{b_G + \alpha + \phi + \delta} \tag{10}$$

E. Market Structure

We consider two market structures for the downstream product: a competitive market, where each participant is small and does not incorporate the effects of his actions on the market price of the downstream product, and a monopoly, where the producer of the downstream product accounts for the effect of his actions on the price of the product.

E.1. Competitive Refinery Sector

In the case of a competitive refinery sector, the optimal quantity and the price of gasoline, as well as the price of crude oil and the crack spread, are given by the solution of the previously described problem:

$$\begin{aligned} Q_G^E &= \min\left(\tfrac{X - P_{I,G} - \Delta + \delta Z + \eta V_Z}{b_G + \alpha + \phi + \delta}, \overline{Q}_G\right) \\ P_G &= X - b_G Q_G^E + P_{I,G} \\ P_C &= \Delta + \alpha Q_G^E \\ P_G - P_C &= X - \Delta - (\alpha + b_G)\, Q_G^E + P_{I,G} \end{aligned} \tag{11}$$

These results suggest, that as long as the marginal production cost increases with the aggregate production, demand shocks are relevant for the downstream product, even in a competitive market. Furthermore, there is a kink in the supply function of the refinery sector: when the demand for the output is too strong, compared to the net global oil supply, the refinery sector works very close to, or at, full capacity, and the production rate is independent of crude oil market situation.

	$Q_G^E < \overline{Q}$	$Q_G^E = \overline{Q}$
P_G^E	$\frac{(\alpha+\phi+\delta)X+b_G(\Delta+P_I-\delta Z-\eta V_Z)}{b_G+\alpha+\phi+\delta}$	$X - b_G\overline{Q}$
P_C^E	$\frac{(b_G+\phi+\delta)\Delta+\alpha(X-P_I+\delta Z+\eta V_Z)}{b_G+\alpha+\phi+\delta}$	$\Delta+\alpha\overline{Q}$
Spread	$\frac{(\phi+\delta)(X-\Delta)-(b_G+\alpha)(\delta Z+\eta V_Z)-P_I}{b_G+\alpha+\phi+\delta}$	$X-\Delta-(\alpha+b_G)\overline{Q}$

Table 2
Equilibrium Values under Competitive Market

The price of the upstream input is determined through the feedback with the demand for the downstream output. However, when the demand for the downstream product is too high, the capital asset operates at maximum capacity, and demand shocks for the downstream product do not influence the price of the upstream input. For example, for the case of crude oil and gasoline, this implies that when refineries operate at capacity, the co-movement between gasoline prices and global crude oil prices is lower than when refineries operate well below capacity. For supporting empirical evidence, see **?** who show that US gasoline demand shocks become insignificant for the global crude oil prices.

E.2. Refinery Sector with Market Power

To provide a benchmark, we calculate the equilibrium quantity for a monopolistic refinery industry. Such an industry considers the effect of its production decisions on the equilibrium prices in the global oil market as well as output market. Therefore, instead of maximizing the total social surplus (TSS), the firm is maximizing the total profit $\pi = P_G(Q)Q - TC(Q)$. Solving this problem gives the equilibrium values for the monopolist industry.

	$Q_G^E < \overline{Q}$	$Q_G^E = \overline{Q}$
P_G^E	$\frac{(2\alpha+b_G+\phi+\delta)X+b_G(\Delta+P_I-\delta Z-\eta V_Z)}{2b_G+2\alpha+\phi+\delta}$	$X - b_G\overline{Q}$
P_C^E	$\frac{(2b_G+\alpha+\phi+\delta)\Delta+\alpha(X-P_I+\delta Z+\eta V_Z)}{2b_G+2\alpha+\phi+\delta}$	$\Delta + \alpha\overline{Q}$
Spread	$\frac{(b_G+\alpha+\phi+\delta)(X-\Delta)-(b_G+\alpha)(\delta Z+\eta V_Z)-P_I}{2b_G+2\alpha+\phi+\delta}$	$X - \Delta - (\alpha + b_G)\overline{Q}$

Table 3
Equilibrium Values under Monopoly Market

$$Q_G^E = \min\left(\frac{X - P_{I,G} - \Delta + \delta Z + \eta V_Z}{2b_G + 2\alpha + (\phi + \delta)}, \overline{Q}_G\right) \tag{12}$$

The price of gasoline, crude oil, and crack spread have the same formulas as in the competitive case in Equation (**??**). The summarized forms are reported in Table **??**

From Equation (**??**) we notice that the optimal quantity of the downstream product produced is lower under a monopoly than under a competitive market structure, unless the capital asset operates at capacity. We also notice that, everything else being equal, under a monopoly the demand for the downstream product (or the price of the upstream input), has to be higher (lower) for the capital asset to operate at capacity.

In general, as long as the operating asset operates at capacity, there are no differences between the two market structures; an indication that, when the capital asset operates at capacity, an inelastic supply side arises in the market for the downstream product. When the capital asset operates below capacity, under a monopoly structure, the price of the downstream output is higher, the price of the upstream input is lower, and the corresponding crack spread is higher than in a competitive market structure.

Parameter	Value
Monthly Interest Rate	0.01
Crude Oil Gasoline Correlation	0.7
Gasoline Demand Mean–Reversion Speed	0.3
Log Gasoline Demand Volatility	0.2
Crude Oil Shocks Mean–Reversion Speed	0.3
Crude Oil Shocks Volatility	4
Maximum Production Rate	10.5
Weight of Recent Production in the Historical Rate	0.3
Cost of Deviation from Historical Rate	4.4
Cost of Producing at Higher Capacity	3.2
Demand Elasticity	-0.95

Table 4
Base–Line Parameters

III. Base Case

We use numerical methods (implemented in MATLAB) to solve for the value function as well as optimal policy function. Table **??** shows the baseline parameters used for the generating the numerical results. We discretized the state space (25 points for each stochastic process) and policy space (200 grid points). Since shocks to crude oil and gasoline market are correlated, we use a MATLAB numerical routine for generating the transition probability matrix over the entire state–space. The value function iteration method is used to find the optimal policies and final value function.

We notice that output level, spreads level, output volatility, and spread volatility follow a regime-switching behavior. When production is well below capacity, production takes place in the competitive region, crack spreads are small, and shocks propagate from the input to the output market and vice–verse. In contrast, when the demand for output is strong or input price is cheap, production is close to capacity, shocks only affect the market where they are originated, input/output prices decouple, and the volatility regime changes significantly. In Figure **??**, the behavior of optimal production in the interior and boundary regions is presented.

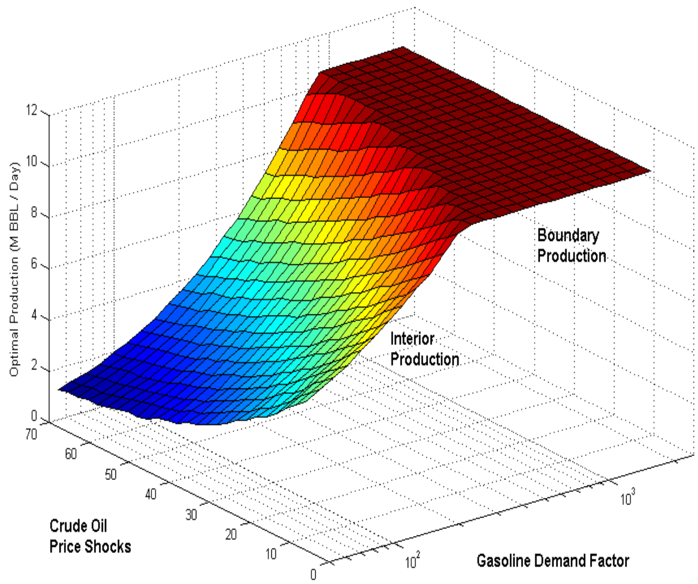

Figure 1. Optimal Production Rates

? evaluate the effect of Katrina on the cost pass through of crude oil on gasoline. Using an ECM model, they show that, in periods before and long after the Katrina, every unit of increase in the price of crude oil results in an increase of almost the same magnitude on gasoline prices. On the other hand, right after Katrina, a 10 cent increase in crude oil price drives up the price of gasoline by 63 cents. They do not provide a theoretical justification for their result while we are able to explain it with our model. In the normal periods, the industry works in a competitive mode where cost pass-through should be close to one. During the hurricane, the effective capacity of industry falls, and the production is in the binding region. Since demand and input shocks are correlated, demand shocks propagate to the price making the seemingly cost pass–through parameter large.

Figure **??** shows spreads versus different values of gasoline demand and equilibrium crude oil prices for a historical production rate of 90%. This graph gives a full description of spreads dynamics in various regions of state–space. We see that spread is convex in the output shocks. When output shock is low, the production is competitive, and demand shock has a weak ef-

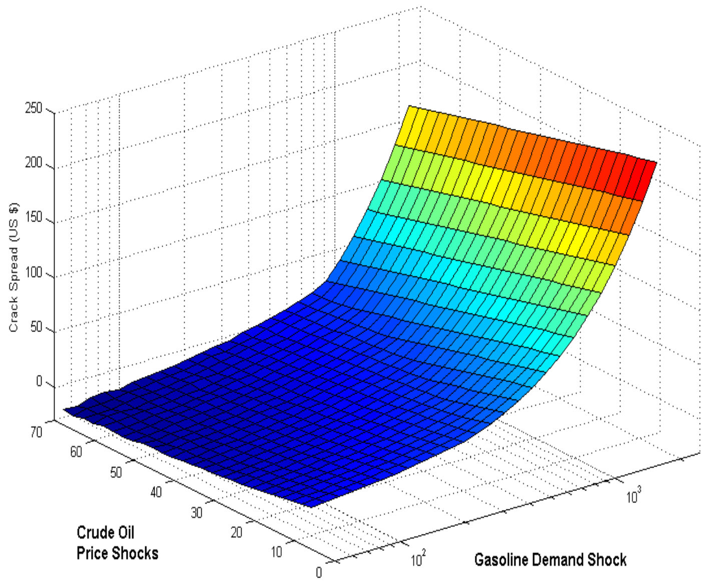

Figure 2. Crack Spreads

fect on output price. However, when demand becomes strong, the production moves to the constrained region, and price increases quickly with demand shocks. Also, in this region the input price has little effect on the output price. When demand is weak input price is the major determinant of output price.

Furthermore, input price and spreads show an interesting dynamic. When input price is low, output production increases to keep the input/output prices in equilibrium. However, there is a limit for this adjustment. When the input price is too low, crack spreads increase quickly because output price can not be lower than a certain amount dictated by refinery capacity constraints. In Aristotle's story about Thales, Thales benefited from this situation.

Observation 1: The correlation between input and output prices is higher in the interior compared to the boundary region

Observation 2: Output prices, spreads, and profits accruing to the capital asset are positively skewed; Spreads and profits are near zero for a significant fraction of time but jump to *a large positive level in certain periods. If the industry is a major source for the input (crude oil), the price of input will be negatively skewed.*

This observation is based on the intuition that, when the capital asset operates below capacity, which is most of the time, profits are close to zero. When the capital asset operates close to capacity, which happens relatively rarely, profits become significant and positive. [6]

Observation 3: The correlation between input and output prices is time–varying and depends on the level of output demand and input supply.

When output demand and supply of input are low, the capital asset operates below capacity, and the price of input and output are highly correlated. When output demand is strong or supply of input is high, the optimal production rate does not respond to the changes in the input price, as long as the input price is not high enough to bring the production back below capacity. Thus, input price shocks do not influence the output price and the correlation between the two decreases significantly.

Observation 4: Spreads can be negative in a dynamic model. An increasing marginal cost decreases the probability of observing negative spreads

Negative spreads are the result of producing higher than static competitive solution. Inspecting the optimal production equation suggests that, when $V_Z > 0$, the dynamic considerations increase the optimal production to a level higher than competitive solution. This happens in two cases: either the level of Z is high and cutting from production is costly, or Z is small but the producer expects reversion to higher levels and gradually increases the production to avoid a sudden increase (which is too costly due to the convexity of adjustment costs).

[6] This distribution of profits has interesting consequences. Policy makers and consumers who observe a high crack spread or high profitability of a converting industry in certain time frame may be tempted to ask for a regulation of the profits of that industry. This partial view ignores the fact that the industry makes profits only for a fraction of its time span. Therefore, reducing the profits (e.g. through taxing) in these unusual moments may destroy the incentives to invest in this industry and reduce the overall welfare.

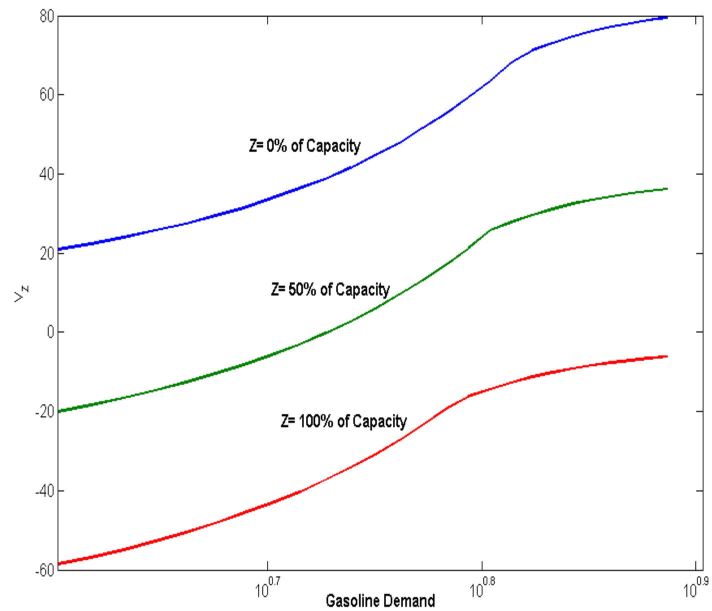

Figure 3. Shadow Value of Average Historical Production

The shadow value of marginal production rate (V_Z) for three different levels of historical production, depicted in the Figure **??**. When demand is weak, $V_Z > 0$ implying that the producer prefers to add to the production rate to get closer to the expected long–run production rate. On the other hand, when demand is strong, the producer expects to stay in this state for some period and tends to increase the historical production rate.

The existence of an increasing marginal cost lowers the average historical production in comparison to the maximum production capacity and as a result, makes it easy to respond to sudden negative demand shocks.

IV. Comparative Statics and Empirical Predictions

The results from our model indicate that there are two regimes, broadly defined, in the relationship between the upstream input and the downstream output in a supply chain. The capital asset operates close to capacity when the price of the input is sufficiently low (or equivalently the supply of the input is sufficiently high), when the demand for the output is sufficiently high, or both. When the capital asset operates close to capacity, we are in a situation of an inelastic supply and quasi–monopoly rents[7]. In that situation the production of the downstream output does not change for small changes in the input supply or the output demand while the spread in price between output and input is large, resulting in large profits for the owner of the capital asset. At the same time, the correlation between the input and output prices drops. On the other hand, when the capital asset operates well below capacity, the spread in price between input and output is small (and possibly negative, depending on the adjustment costs), the owner of the capital asset makes little or no profit, and the price of the input and output are highly correlated.

Given this intuition, we offer a few propositions for the steady–state case (in which $V_Z = 0$ and it is a proxy for long–run behavior of the system) that can be empirically tested, based on how changes in the values of certain parameters make it more likely to enter the regime where the capital asset operates at capacity.

Proposition IV.1. *If the long–run optimal production is below the capacity constraints, higher volatility of demand for the downstream output, and the supply of the upstream input increase the frequency of hitting the boundary region.*

Proposition IV.2. *If the long–run optimal production rate is below the capacity constraints, expected spreads between input and output prices increase with the volatility of output and input. The effect is stronger when the level of mean demand for the output is high, or when the price of the input is high.*

Proposition IV.3. *Higher correlation between input and output shocks weakens the positive effect of volatility on the frequency of hitting the boundary and on expected spreads.*

[7] We use quasi–monopoly rent instead of *monopoly rent* because the later requires collusions and coordination among players of the industry

A joint formal proof is given in Appendix **??**. The intuition behind these observations is that, everything else being equal, higher volatility makes it more likely to reach the regime in which the capital asset operates at capacity. This is the region where quasi–monopoly profits accrue to the capital asset. However, the volatility effect is not monotonic. If the long–run production rate is above the capacity constraints, an increase in the volatility may decrease asset value because it will push production back to the competitive region. In summary, the owners of assets prefer unexpected volatility shocks and declined correlation in input and output markets.

From Equation (**??**), the crack spread, and the optimal quantity of output produced, are expressed as a weighted difference between the demand for the output and the supply of the input. As the correlation between these two quantities increases, their difference becomes less volatile (resembling the portfolio theory in the case of two asset with negative correlations), leading to a few instances in which the optimal production amount is close to capacity. This result suggests that, everything else equal, the conversion technology earns higher profits in industries where the input and output prices are volatile but not correlated with each other.

Proposition IV.4. *Variance of spreads is smaller in the interior.*

Proof. The variance of spreads in the interior and the boundary are:

$$\begin{cases} \mathrm{Var}(\mathrm{CS})_{\mathrm{Interior}} = (\frac{\phi+\delta}{b_G+\alpha+\phi+\delta})^2(\mathrm{Var}(X)+\mathrm{Var}(\Delta)) \\ \mathrm{Var}(\mathrm{CS})_{\mathrm{Boundary}} = \mathrm{Var}(X)+\mathrm{Var}(\Delta) \end{cases} \tag{13}$$

Since $0 \leq (\frac{\phi+\delta}{b_G+\alpha+\phi+\delta})^2 < 1, \Rightarrow \mathrm{Var}(\mathrm{CS})_{\mathrm{Interior}} < \mathrm{Var}(\mathrm{CS})_{\mathrm{Boundary}}$

□

Proposition IV.5. *Spreads are smaller in the competitive market.*

Proof. Spreads are the same in the boundary region for both market structures. In the interior they are given by:

$$\begin{cases} CS_{\text{Competitive}} = \frac{(\phi+\delta)(X-\Delta)+(b_G-\alpha)(\delta Z+\eta V_Z)-P_I}{b_G+\alpha+\phi+\delta} \\ CS_{\text{Monopoly}} = \frac{(b_G+\alpha+\phi+\delta)(X-\Delta)+(b_G-\alpha)(\delta Z+\eta V_Z)-P_I}{2b_G+2\alpha+\phi+\delta} \end{cases} \tag{14}$$

Since $\frac{\phi+\delta}{b_G+\alpha+\phi+\delta} < \frac{b_G+\alpha+\phi+\delta}{2b_G+2\alpha+\phi+\delta} \Rightarrow CS_{\text{Competitive}} < CS_{\text{Monopoly}}$

□

Proposition IV.6. *Spreads are less volatile in the competitive market.*

Proof. The proof is the same of the proof for the previous Proposition, by taking the variance.

□

Proposition IV.7. *Spreads are smaller when the elasticity of demand is higher*

Proof. Consider the relationship for spreads in the competitive market: $\frac{(\phi+\delta)(X-\Delta)+(b_G-\alpha)(\delta Z+\eta V_Z)-P_I}{b_G+\alpha+\phi+\delta}$. Observe that in the steady–state, $\frac{\partial \text{Spread}}{\partial b_G} < 0$. The same argument applies for the monopoly market. □

Proposition IV.8. *The variance of gasoline price is regime–dependent. There is no a-priori ordering of variance between two regions.*

Proof. The variance of output price in the interior and in the boundary regions are is given respectively by:

$$\begin{cases} \text{Var}(P_G) = \frac{(\phi+\alpha+\delta)^2\text{Var}(X)+(b_G)^2\text{Var}(\Delta)+b_G(\phi+\alpha+\delta)\text{Cov}(X,\Delta)}{(b_G+\alpha+\phi+\delta)^2} \\ \text{Var}(P_G) = \text{Var}(X) \end{cases} \tag{15}$$

These equations suggest that, in the interior, the variance is a linear combination of the variance of input price and output demand shocks. In the boundary, it is just the function of the variance of the demand shocks. Therefore, the variance changes between two regions. However, depending on the value of coefficients, it can be larger or smaller in the interior versus the boundary region.

□

V. Empirical Evidence

In the previous sections, we have provided a model for expected spreads and refinery value. The model implies that expected profits increase in input and output volatilities, as well as output price, and decrease in input price. To test the empirical relevance of our model, we use the stock market and energy commodity price data. Changes in crude oil and gasoline level and volatilities are major drivers of refiners' "cash–flow" shocks. Other papers (e.g. **?**) show that individual stock returns are mostly driven by idiosyncratic cash–flow shocks, and therefore, we interpret our results as a support for the cash–flow generation model, rather than a risk–return compensation or discount–rate effect.

A. Data

We employ daily stock prices of four major refining companies (reported in Table **??**) as a proxy for representative refinery value. Finding appropriate stocks to be included in the regression is a challenge because there are few publicly traded companies that have refining as their dominant business line. The majority of large and well–known energy companies (e.g. BP, Chevron, Exxon Mobil) are not good choices for this purpose because they are active in both upstream and downstream markets, and thus their cash–flow has parallel positive and negative exposures to crude oil price movements. The best stock to represent a pure refining

Company	Ticker	Country
Valero	VLO	US
Sunoco	SUN	US
Tesoro	TSO	US
Reliance	RIL	India

Table 5
List of Companies

industry is Valero Energy Corporation, one of the largest refiners in the US that is mainly focused on the refining business. We have also included three other companies for robustness checks: *Sunoco* and *Tesoro* are both active in the US (though smaller and more diversified than Valero), while, *Reliance Group* owns world's largest refinery operating in India, though traded on the US stock market. Crude oil and gasoline spot prices in daily frequencies are collected from the EIA website. Monthly futures prices of gasoline and crude oil are obtained from Bloomberg. We use the S&P500 index as market factor and the IRX index (13–week Treasury bill) as a proxy for risk–free rate. To control for non–CAPM risk factors, we include monthly SMB and HML series obtained from Kenneth French's website. Finally, as control variables for omitted variable bias, we include 5–year treasury yield, average retail price of gasoline, US refinery production, US refinery operable capacity, and US gasoline import. The data range from 1986/06 until 2010/10 (6197 daily observations). The return data are converted to monthly frequency to avoid problems of high-frequency return estimations (final 292 monthly observations for each firm). We estimate the monthly volatility of crude oil and gasoline prices by calculating the standard deviation of daily prices inside of that given month. Since the resulting time–series are not stationary, we use the first difference of all variables in the regression and do the standard unit–root tests on the differenced values. We also, make sure that the residuals are not serially correlated. For return variables, we use simple monthly returns.

B. Econometrics Set–Up

We employ the following basic regression, where r_i is the monthly excess return on the stock of the refining company, r_m the excess return on market factor proxied by the S&P500 index, *Spread* the ratio of gasoline to crude oil prices, *Gas Volatility* and *Crude Volatility* are the intra–month volatility derived from daily spot prices, Spreads_T are the spreads for a T months maturity, r_O the excess return on crude oil, Ψ the vector of controls and $d()$ is the difference operator.

The model implies that the higher the capacity utilization of the converting industry is, the higher the spreads (and profits) will be. Therefore, we also include the ratio of imported to domestically produced gasoline to measure the tightness in the production process:

$$
\begin{aligned}
r_i = \beta_0 + \beta_1 r_m + \beta_2 \text{d(Spread)} + \beta_3 \text{d(Gas Volatility)} + \beta_4 \text{d(Crude Volatility)} + \\
\overline{\beta}_F d(\text{Spread}_T) + \beta_O r_O + \beta_C \Psi + \beta_L \frac{\text{Import}}{\text{Production}} + \varepsilon
\end{aligned}
\tag{16}
$$

The results from the standard OLS regression for *Valero* are reported in Table **??**. Our theoretical model predicts that the value of refinery increases with demand shocks and decreases with input prices (two effects proxied by spot and forward spreads). Theory also suggests that the volatility of both input and output markets increase the refinery value. Positive, significant coefficients of crude volatility support the theory. Because of a break in 2008, we include a dummy to estimate volatility effect outside of this year. It can be read from Table **??** that adding the level and volatility of underlying factors increases R^2 from 10% to 32%, and can explain totaly 22% of variation in stock prices. In the basic regressions where the levels of crude oil and gasoline are used, one can notice that the coefficients on crude oil and gasoline price levels are close but with opposite signs. This suggests that the crack spread (ratio of output to input prices) is a better measure for refinery value since it eliminates the basic level effect, too. Forward contracts of commodities contain information about the long–term trends, whereas spot prices may vary with local (short–term) conditions. Hence, we use both the spot and forward spreads (instead of individual price levels) in the reported regressions.

Several recent studies (e.g. **?**, **?**, **?**) suggest that the crude oil price is a new risk–factor in the global economy. However, this is not the same effect we find for our cash–flow factors. Refiners are consumers of crude oil. If oil price volatility was merely a risk–factor for refining industry, then a positive return to oil price and a decrease in the level of crude oil volatility should increase the contemporaneous stock return of the company since lower crude volatility lowers the risk premium and discount rate. On the contrary, we find that crude oil volatility *increases* and the level of crude oil *decreases* stock value (negative return) and this is completely opposite to the risk–factor results reported previously. Based on our results, we interpret crude oil price and volatility as "cash–flow" factors for refining firms, consistent with our theoretical model.

To conduct some robustness checks we repeat the final regression for the other three companies and report the results in Table **??**. The results are not far from expectations. The level factors are significant for US–based firms. However, the crude oil volatility effect is not significant. One reason might be the mix of different business lines in these two companies. Furthermore, Reliance is a foreign company traded in the US stock market. Hence, the fundamental cash–flow factors relevant for the US market need not be relevant to its stock price.

Since stock market data are forward looking, any new information should immediately be incorporated into prices by rational agents in the market. To be consistent with this behavior, we should not see any predictive power for the lagged variables. We test different lags of all variables separately and none of them turn out to be significant.

To make sure that our results are not the effect of a missing variable bias (where the crude oil and gasoline are proxying for the missing variable), we repeat the major regression for a set of energy and non–energy companies. As reported in Table **??** none of the refinery fundamental factors are significant for non-energy companies. Energy companies may have

Variable	Dependent Variable: Excess Return of Valero Stock							
SP500	0.90^{***}	0.89^{***}	0.90^{***}	0.89^{***}	0.91^{***}	0.90^{***}	0.95^{***}	0.95^{***}
	(5.31)	(5.68)	(5.72)	(5.67)	(5.79)	(5.93)	(6.28)	(6.29)
Gasoline Spot		0.01^{***}						
		(3.39)						
Crude Spot		-0.0176^{***}						
		(-3.64)						
Gasoline 3–Month		0.009^{***}						
		(2.60)						
Crude 3–Month		-0.006						
		(-1.46)						
Spreads			0.74^{***}	0.57^{***}	0.57^{***}	0.68^{***}	0.70^{***}	0.68^{***}
			(6.08)	(4.34)	(4.38)	(5.25)	(5.54)	(5.38)
Spreads_3				0.45^{***}	0.45^{***}	0.42^{***}	0.41^{***}	0.42^{***}
				(4.35)	(3.32)	(3.24)	(3.14)	(3.24)
Crude Volatility		0.011			0.011	0.007	0.022^{***}	0.02^{**}
		(1.59)			(1.45)	(1.16)	(2.46)	(2.25)
D_{2008} * Crude Volatility		-0.02^{*}					-0.32^{**}	-0.030^{**}
		(-1.81)					(-2.29)	(-2.30)
Crude Excess Return		0.41^{***}				0.30^{***}	0.45^{***}	0.46^{***}
		(2.93)				(4.26)	(5.30)	(5.35)
D_{2008} * Crude Excess Return		-0.48					-0.55^{**}	$-.56^{***}$
		(-1.36)					(-2.28)	(-2.33)
Gas Import Production Ratio								0.66^{*}
								(1.25)
SMB	0.007^{***}	0.006^{***}	0.006^{***}	0.006^{***}	0.006^{***}	0.006^{*}	0.007^{*}	0.007^{*}
	(2.88)	(2.77)	(2.64)	(2.52)	(2.65)	(2.88)	(3.21)	(3.21)
HML	0.00	0.00	0.00	0.00	0.00	0.002	0.00	0.00
	(0.29)	(0.29)	(0.17)	(0.46)	(0.80)	(0.39)	(0.45)	(0.47)
R^2	10.9%	28.0%	21.0%	24.0%	24.5%	30.0 %	32.8%	33.1%

Table 6: Return Factors of Refinery Industry for Valero Co

Variable	Dependent Variable: Excess Return of Stock		
	Sunoco	Tesoro	Reliance
SP500	0.73^{***}	1.04^{***}	0.43^{**}
	(6.79)	(6.04)	(2.06)
Spreads	0.24^{***}	0.42^{***}	-0.10
	(2.75)	(2.98)	(-.55)
Spreads_3	0.21^{**}	0.45^{***}	0.089
	(2.32)	(3.06)	(0.52)
Crude Volatility	0.003	0.01	0.01
	(0.61)	(1.24)	(1.18)
D_{2008} * Crude Volatility	-0.02^{***}	-0.007^{***}	-0.007
	(-2.79)	(-2.34)	(-0.56)
Crude Excess Return	0.05	0.25^{***}	0.17
	(0.89)	(2.57)	(1.49)
D_{2008} * Crude Excess Return	-0.53^{***}	-0.77^{***}	-0.007
	(-3.10)	(-2.83)	(-0.03)
SMB	0.00	0.009^{***}	0.006^{**}
	(0.12)	(3.58)	(2.28)
HML	0.00	0.006^{***}	0.005^{***}
	(0.12)	(3.37)	(2.91)
R^2	21.9%	24.2%	18.1%

Table 7
Return Factors of Refinery Industry

some exposure, and it is therefore expected that some of the factors are significant for them, but not as cleanly and consistently as for Valero.

VI. Model Calibration and Estimation of Parameters

We estimate the parameters of demand (long–run mean and volatility) and the parameters of the production function (convexity cost and adjustment costs) of the structural model built in the section **??**. Usually, the demand parameters of product markets (demand elasticity in particular) are estimated using an instrumental variables approach or benefiting from exogenous shocks to supply, in order to overcome the issue of endogeneity of the price and quantity. One can see that a simple regression of price on quantity will provide biased estimations because quantity is an endogenous variable. However, our estimation is consistent in this respect. Instead of using a reduced–form price model, which ignores the endogenous responses of supply side, we use a full structural model, with two separate equations for price and quantity. The model generates endogenous values of quantity and price simultaneously. Therefore, we are able to identify the effect of shocks coming from demand and supply side, separately. By matching the outputs of our model to certain moments of the data, we are able to recover unobservable parameters of the model.

We use a combination of the Kalman filter (to recover seasonality features) and the Simulated Method of Moments (SMM) techniques. The SMM approach assumes that the model–generated sample and the empirical data are drawn from the same distribution. However, the parameters of the distribution are unknown and should be estimated. By matching certain moments of model's output with the empirical moments of the data, one is able to estimate the unknown parameters.

Our estimation contributes to the literature in two aspects. First, unlike many papers that use reduced–form models and report an average value, we treat demand parameter as a stochastic process, then estimate its dynamics using a structural model. Second, in addition

Variable	Dependent Variable: Excess Return of Stock					
	BP	Chevron	Exxon Mobil	Conoco	GE	Walmart
SP500	0.75^{***}	0.69^{***}	0.56^{***}	0.86^{***}	1.25^{***}	0.85^{***}
	(10)	(11.18)	(10.8)	(10.8)	(19.96)	(11.14)
Crude	-0.01^{***}	-0.009^{***}	-0.007^{***}	-0.009^{***}	0.002	0.001
	(-4.76)	(-4.92)	(-4.43)	(-3.74)	(1.57)	(0.58)
Gasoline	0.001	0.001	0.001^{*}	0.003^{**}	-0.001	-0.00
	(1.21)	(1.182)	(1.76)	(2.15)	(-1.58)	(-0.40)
$Crude_3$	0.009^{***}	-0.007^{***}	0.005^{***}	0.008^{***}		
	(3.72)	(3.85)	(3.08)	(3.65)		
$Gasoline_3$	-0.001	0.00	-0.001	0.00		
	(-0.62)	(-0.53)	(-1.20)	(-.39)		
Crude Volatility	0.008^{**}	0.00	-0.002	0.002	-0.001	0.001
	(2.31)	(0.14)	(-1.12)	(0.61)	(-0.44)	(0.32)
Crude Excess Return	0.35^{***}	0.25^{***}	0.19^{***}	0.17^{***}	-0.05	-0.13
	(5.24)	(4.56)	(4.19)	(2.42)	(-1.04)	(-1.87)
Gas Import Production Ratio	-0.09	-0.04	0.15	-0.06	-0.21	0.32
	(-0.34)	(-0.18)	(0.85)	(-0.22)	(-0.98)	(1.19)
SMB	-0.001	0.003^{***}	-0.002^{***}	0.001	-0.007^{**}	-0.003^{***}
	(-1.37)	(3.32)	(-3.67)	(-0.95)	(-1.89)	(-2.94)
HML	0.001	0.001^{*}	0.000	0.002^{***}	0.001	-0.002^{***}
	(1.50)	(1.69)	(0.85)	(2.90)	(1.61)	(-2.64)
R^2	39.9%	41.9%	36.2%	42.8%	59.5%	32.5%

Table 8: Return Factors of Refinery Industry Applied to Other Companies

to estimating the dynamics under the physical measure (using spot prices), we calibrate the model under risk–neutral measure. For this purpose, we use futures prices of gasoline to find demand dynamics under the $\mathbb{Q}$–measure. This approach and the resulting parameter values allow us to use our model for valuation purposes.

A. Data

Our data consists of the monthly data from Jan 1983 to Oct 2010. From the Energy Information Administration (EIA) website, we obtain the following data items:

- Average US wholesale gasoline price
- Average cost of crude oil input to U.S refineries
- Aggregate gasoline production
- Monthly import of gasoline

Furthermore, we have collected from Bloomberg the futures prices of the crude oil (ticker CL) and gasoline (ticker HU (1987–2005) and later XB (2005–2010)) up to 12–month maturity for the period of June 1987 to Oct 2010. Finally, we use the CPI index (for urban consumers) from the Bureau of Labor Statistics to exclude the effect of inflation on crude oil and gasoline prices.

B. Set–up

In the section **??**, we used a linear demand function for illustrative purposes. However, for the estimation, we are more flexible and examine different functional forms for gasoline demand. Finally, we decide to choose a log–log (constant elasticity) demand function, given by the equation **??**, and solve the dynamic model using this demand function. Choosing a constant elasticity demand comes with a cost: there is no analytical solution for the first order conditions characterizing the optimal production. However, since we use numerical techniques to solve the dynamic model, this is not a big issue. In return, we are to make a much better fit of the model's output to the data.

$$\ln(P_G) = x + \ln(f_g) + \gamma \ln(q_G^S) \tag{17}$$

Here, $x = \ln(X)$ is the log demand factor, f_g is the time–invariant monthly coefficient (a 12 x 1 vector), q_G^S the total supply of gasoline in the US market (sum of refinery production and imports), and γ is the dynamic elasticity parameter. Other studies have already reported average values for $\gamma \in [-0.25, -1.7]$. Therefore, we can back check whether our estimated γ is in accordance with other studies. The dynamic of x is given by $dx = \mu_x(\bar{x} - x)dt + \sigma_x dW_x$ and unknown parameters $\{\mu_x, \bar{x}, \sigma_x\}$ have to be estimated. Since the dynamic of assets is different under the physical (or $\mathbb{P}$) and risk-neutral (or $\mathbb{Q}$) measures, we need to recover the long–run mean of the demand factor under both measures. We do this separately in the second and third steps of our estimation.

To make it comparable with usual econometrics models of gasoline demand estimation, we can write the general form used in these studies:

$$\ln Q = \beta_0 + \beta_1 \ln P + \beta_2 \ln Y \tag{18}$$

Where $\ln Y$ is the log of income, and β_1 and β_2 are price and income elasticities, respectively. A conversion of this equation to our formulation will give:

$$\ln P = -\frac{\beta_0}{\beta_1} + \frac{1}{\beta_1} \ln Q - \frac{\beta_2}{\beta_1} \ln Y \tag{19}$$

Period	$\overline{P_C}$	σ_{P_C}
1983-2005	23.8	1.61
2005-2010	66.5	6.7

Table 9
Estimated Parameters of Crude Oil Price Dynamics

Therefore, $-(\frac{\beta_0}{\beta_1} + \frac{\beta_2}{\beta_1} \ln Y)$ is equivalent to our x term, which we treat as stochastic process.

We also remind the reader about the cost function of refiners in which we replace Δ with more observable variable P_C:

$$TC(q, P_C, Z) = (P_C + P_I)q + \frac{\phi q^2}{2} + \delta_Z (q - Z)^2 \tag{20}$$

We need to estimate $\theta_2 = \{\phi, \delta_Z\}$. At the end we have 18 major parameters to be estimated: $\theta = \left\{\gamma, F_g, \bar{x}^P, \bar{x}^Q, \sigma_x, \phi, \delta_Z\right\}$. We do this in three steps. In the first step, the 12×1 vector of seasonality coefficients $\mathbf{F}_g$ will be estimated. In the second step, we will estimate $\theta = \left\{\gamma, \bar{x}^P, \sigma_x, \phi, \delta_Z\right\}$, using spot prices. Finally, the risk-neutral long–run mean of demand factor, $\bar{x}^Q$, will be estimated in the third step using futures prices.

Notice that, in the original model, crude oil price is endogenous and given by:

$$P_C(q) = \Delta + \alpha q \tag{21}$$

In order to reduce the dimension of estimation, we assume $\alpha = 0$, and use the crude oil price directly instead of estimating the latent Δ factor. Given the existing empirical estimations on the effect of U.S. gasoline on global crude oil price, this assumption is reasonable. Assuming a mean–reverting process for crude oil price, we estimate its parameters from the observable data and report it in Table **??**

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
-0.04	-0.04	-0.03	0.04	0.05	0.05	0.04	0.03	0.02	-0.03	-0.04	-0.05

Table 10
Monthly Factors, Estimated in the First Step

C. Estimation Strategy

Our strategy to estimate unknown parameters consists of three steps. In the first step, we use a standard Kalman filtering method to recover the monthly factors from the futures prices. In the second step, we use the dynamic model to find values for long–run demand level, the speed of mean–reversion, demand shock volatility, the convexity cost and the cost of deviation from historical rates. In the final stage, we use the calibrated model to estimate the dynamics of demand factor under the risk–neutral measure.

C.1. Step 1: Seasonality Factors

The shape of crude and gasoline futures (see Figure **??** for a sample of 10 different starting dates) suggests that, unlike crude oil, gasoline futures have a strong seasonal element. This justifies the seasonality factor $F_G = [F_{G,T}], T \in \{1, ..., 12\}$ to capture the seasonality effect of the month. One can estimate the seasonality factors by simply running a dummy regression over the spot prices. However, futures prices provide a much richer and less noisy information set to estimate the monthly seasonality factors. We employ a standard Kalman filter exercise to extract monthly factors of gasoline. The estimation results are reported in Table **??**.

Estimated monthly factors suggest that the *high demand* seasons include April through September. Although the major driving season starts in June, the market stores gasoline produced in April and May to respond a high demand shock in June. That is why we see some seasonality factor in April and May, too.

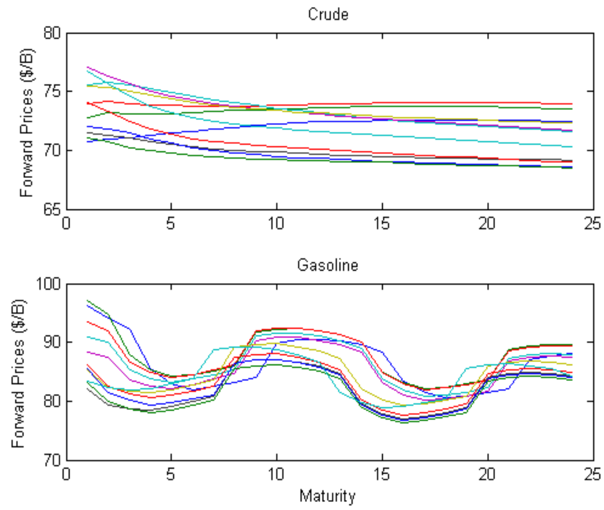

Figure 4. Samples of Crude Oil and Gasoline Futures

C.2. Step 2: Production and Demand Parameters under ths Physical Measure

At any time point t, we observe spot input crude oil price $P_{C,t}$ and calculate the price difference $dP_{C,t-1} = P_{C,t} - P_{C,t-1}$. Given that the crude oil price has a mean–reverting dynamics, the realized exogenous shock to the crude oil market can be estimated by:

$$dW_{C,t-1} = \frac{dP_{C,t-1} - \mu_{P_C}(\overline{P} - P_{C,t-1}))}{\sigma_{P_C}} \tag{22}$$

Shocks to crude oil and gasoline markets are correlated. Hence, we generate the shock to the gasoline market through the following equation:

$$dW_{G,t-1} = \rho dW_{C,t-1} + \sqrt{1-\rho^2}dW \tag{23}$$

Here ρ is the correlation between crude oil and gasoline market shocks, and dW is the standard Gaussian shock. This way, we capture the correlation feature of our model. The updated value of the gasoline demand factor is:

$$X_{G,t} = X_{G,t-1} + \mu(\overline{X} - X_{G,t-1}) + dW_{G,t-1} \tag{24}$$

The endogenous optimal production q_t is given via the first order conditions of the dynamic model by plugging $\{x_{G,t}, P_{C,t}, Z_t\}$ into the optimal policy function of the numerically solved model. Since we use a discrete state–space for the numerical calculation, the optimal quantity is found by interpolating the optimal policies at points around $P_{C,t}$ and $P_{G,t}$.

Having $x_{G,t}$ and q_t generated and Import_t collected, we are able to generate gasoline price , $P_{G,t}$, from the model. Repeating this procedure for $t \in [1:T]$, provides us with a $2 \times T$ vector of model–generated gasoline prices and gasoline production for the entire time period. We iterate the procedure 300 times and calculate the averaged values of the following moments:

- Gasoline price
  - Mean
  - Variance
  - Skewness
- Gasoline production
  - Mean
  - Variance

We call the vector of model–generated outcomes by $\chi(\theta)$ and the moment function by $h(.) : \mathbb{R}^2 \to \mathbb{R}^5$, and then produce the averaged moments of model's output:

$$M(\theta)_{\text{Model}} = \frac{1}{K} \sum_{1}^{K} \left(\frac{1}{T} \sum_{t=1}^{t=T} h(\chi_t(\theta)) \right) \tag{25}$$

Here, M_{Model} is the vector of model–generated moments , K is the number of iterations and T is the length of observations. We, also, calculate the counterpart empirical values of all these moments for the empirical data, M_{Data}. We minimize the following objective:

Parameter	Value
γ	-0.9857
$\bar{x}$	5.77
σ_x	0.2
δ_Z	4.4
ϕ	3.35

Table 11
Demand and Production Function Parameters, Estimated under the Physical Measure

$$\mathbb{E}(\theta) = (M_{\text{Data}} - M_{\text{Model}}(\theta))\mathbb{W}(M_{\text{Data}} - M_{\text{Model}}(\theta)) \tag{26}$$

$\mathbb{W}$ is a 5×5 positive-definite weighting matrix, calculated in a two–stage procedure. Finally, there are 5 moment conditions, and we need to estimate 5 parameter values. Therefore, our model is just–identified from an estimation point of view. We use numerical methods in Matlab, to solve the dynamic model and minimize the objective function.

$$\hat{\theta} = \arg\min \left\{ \mathbb{E} = (M_{\text{Data}} - M_{\text{Model}}(\theta))' \mathbb{W}(M_{\text{Data}} - M_{\text{Model}}(\theta)) \right\} \tag{27}$$

The minimization of the objective function results in values for unknown parameters θ, which are reported in Table **??**. What we find for the elasticity parameter stays in the range other researchers have reported. We also find positive values for producing at higher capacity, as well as deviating from the historical production rate.

Figure **??** shows a sample of model–generated time–series for gasoline price and quantity versus actual data for the time period of 1986–2005. Obviously, the model does a better job in predicting production levels in the later than the initial periods. One possible explanation would be that we have assumed a constant long–run mean for the demand parameter and elasticity, while, in reality the elasticity has decreased over time.

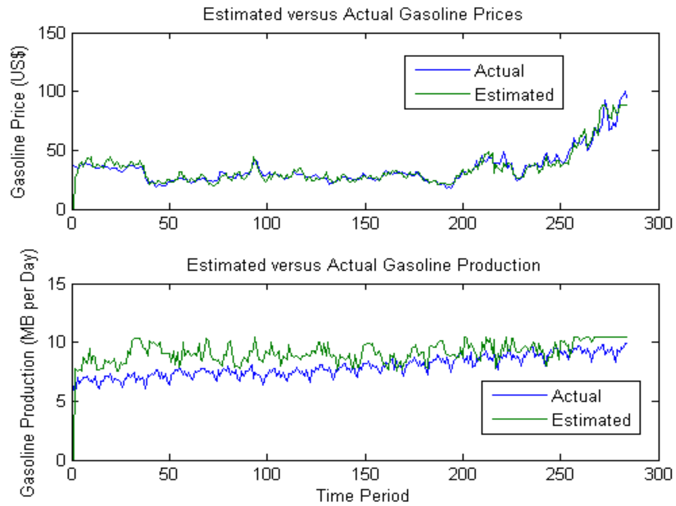

Figure 5. Model Performance versus Empirical Values

C.3. Step 3: Demand Parameters under ths Risk-neutral Measure

Gasoline futures prices are rich sources of information for estimating dynamic parameters under the risk–neutral measure. Since all contracts in futures markets are priced under the risk–neutral measure, a model calibration using these data will provide us with risk–neutral parameters. These parameters can be used for pricing various assets (e.g. spread options) whose underlying asset is gasoline price.

Our estimation strategy is as follows. At each calendar time t, we observe the futures prices of crude oil and gasoline, up to the maturity of 12 months. We use the model to generate futures prices of gasoline up to the same maturity, using crude oil futures as an input to the model. We generate future prices by producing 200 price samples for each maturity and then averaging the prices for each maturity.

To find the expectation of the latent demand factor for each maturity, we use the following standard formulation, also used by other authors (e.g. **?** and **?**). Denoted the log of the

Period	$\bar{x}$	μ_Q
1986–2005	5.15	0.19

Table 12
Demand Estimation under the Risk–neutral Measure

stochastic demand parameter at time t by $x(t)$, the dynamics of this process under the Q–measure will be:

$$dx = \mu_Q(\bar{x}_Q - x)dt + \sigma_Q dW \tag{28}$$

Taking the expectation gives the expected value of x at any future maturity time T.

$$E(x(t,T)) = x(t)e^{-\mu_Q(T-t)} + \bar{x}_Q(1 - e^{-\mu_Q(T-t)}) \tag{29}$$

.

Since the volatility is measure–invariant, we assume the volatility term σ_Q to be the same as σ_P. Therefore, we have to estimate $\left\{\mu_x^Q, \bar{x}^Q\right\}$ from the model.

To estimate these two parameters, we use the production function parameters from the previous step, and match the first moments of gasoline futures prices to model–generated moments (12 moment conditions), accounting for the seasonality effect. The estimation results are reported in Table **??**.

As expected, the long–run mean of demand process under the risk–neutral measure is smaller than under the physical measure.

D. Discussion of Results

Since we do not model the storage in our model, our results might be biased. In the presence of storage, the optimal production will be less sensitive to demand shocks and more sensitive to input shocks because the refiners can produce in low–demand periods and store the gasoline to sell in periods with higher demand, and possibly with more expensive input.

Seasonal shut–downs (for maintenance purposes) are another feature we do not model. Refiners need to stop operation and do necessary maintenance operations to lower the probability of break–downs. The planned shut-downs generate seasonalities in the production rate, which we don't capture. [8]

VII. Conclusions

We have presented an equilibrium model for the price of the output in a supply chain, given a stochastic process for the supply of the input, the demand of the output, and a capital asset that transforms the input into the output and faces operational constraints. We calibrated our model to the case of a refinery turning crude oil into gasoline.

One application of the model is to determine the profit generating process of such industries, its asset price implications, and the optimal hedging policy a refinery. For example, a consequence of our model is that, if demand for gasoline is weak, input and output prices follow almost the same path with a fixed crack spread. In this case, the refinery does not need to hedge at both sides, because changes in the price of the input are offset by changes in the price of the output, leaving a margin for the refinery. On the other hand, if demand for gasoline is strong, output and input prices are no longer highly correlated, resulting in a more volatile crack spread. Using the dynamics of the crack spread from our model, the refinery can decide whether a hedge of input or output is necessary and then calculate the optimal hedge policy.

Another application of our model is to use the estimated cash flows under the risk-neutral measure to estimate the value of the financial contracts such as futures or options on gasoline, or even the value of the refinery itself. Beyond simply valuing the refinery we can also derive

[8] For a detailed discussion of planned shut–downs, see **?**

the dynamics of the covariance of the value of the refinery with the market portfolio — beta — given information on the covariance of the stochastic processes of the supply of crude oil, and the demand for gasoline with the market portfolio. We leave these applications for future research.

Appendix

A. Empirical Evidence

In this section, we provide empirical facts about the behavior of crack spreads, inter–month correlation of crude oil and gasoline prices and intra–month gasoline price volatility. All data are obtained from EIA website.

Figure **??** shows a time-series of crack spreads for 1992–2010. In Figure **??** the intra–month correlation between crude and gasoline is presented, showing a volatile regime of correlation which is becoming even negative in some months.

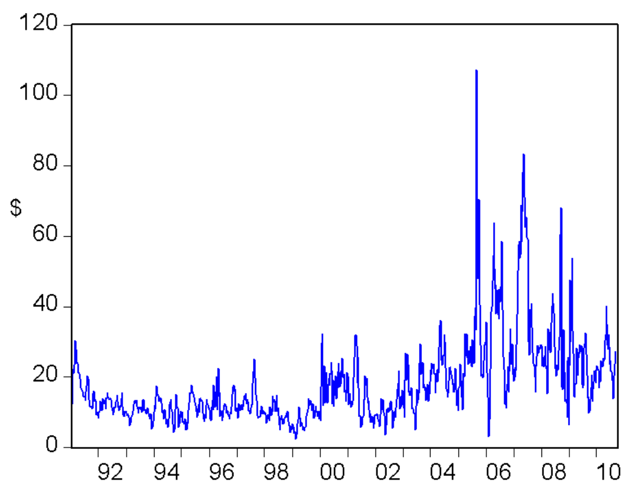

Figure 6. Crack Spreads, 1992–2010

The average yields of various refined products are presented in Figure **??**. These numbers suggests that the sum of gasoline, distillate and jet fuel yields, account for approximately 80% of the output of a typical refinery.

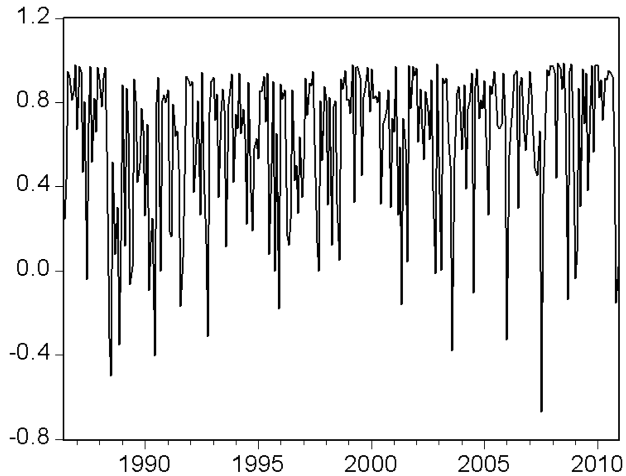

Figure 7. Intra–Month Correlation of Crude Oil and Gasoline Prices

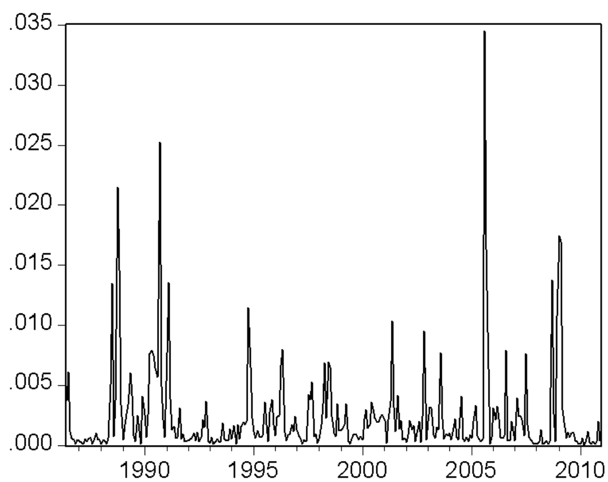

Figure 8. Monthly Volatility of Gasoline Prices

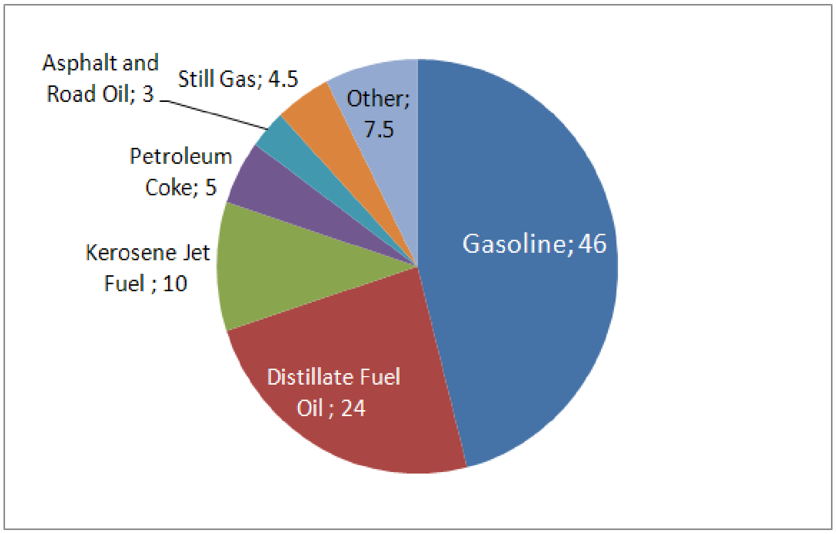

Figure 9. Average Yields of Refined Products (%): 1993-2010

B. Effect of Volatility on Spreads and Asset Value

If $\delta =$, the optimal production is given by $Q_G^E = \min(\frac{X-\Delta}{b_G+\alpha+\phi}, \overline{Q})$ (for simplicity we assume $P_I = 0$). Assume the current production is in the interior. Using the Bayes' rule, we can calculate the expected spread and profit for the next period:

$$\mathbb{E}(\text{Spread}_{t+1}) = P(Q^E < \overline{Q})\mathbb{E}(\text{Spread}_{t+1}|Q^E < \overline{Q}) + P(Q^E = \overline{Q})\mathbb{E}(\text{Spread}_{t+1}|Q^E = \overline{Q}) \quad (30)$$

$$\mathbb{E}(\pi_{t+1}) = P(Q^E < \overline{Q})\mathbb{E}(\pi_{t+1}|Q^E < \overline{Q}) + P(Q^E = \overline{Q})\mathbb{E}(\pi_{t+1}|Q^E = \overline{Q}) \quad (31)$$

If the current values for gasoline and crude oil factors are X_t and Δ_t, the next period values will have the following distributions:

$$X_{t+1} \sim \mathbb{N}(\mu_X \bar{X} + (1-\mu_X)X, \sigma_X^2) \quad (32)$$

$$\Delta_{t+1} \sim \mathbb{N}(\mu_\Delta \bar{\Delta} + (1-\mu_\Delta)\Delta, \sigma_\Delta^2) \tag{33}$$

Furthermore, these two variables are correlated with each other with a correlation coefficient of ρ. The production will be in the interior as long as

$$X - \Delta < (b_G + \alpha + \phi)\overline{Q} \tag{34}$$

Define $Z = X - \Delta$ and observe that Z is normally distributed with

$$Z \sim \mathbb{N}(\mu_X \bar{X} + (1-\mu_X)X - \mu_\Delta \bar{\Delta} + (1-\mu_\Delta)\Delta, \sqrt{\sigma_X^2 + \sigma_\Delta^2 - \rho\sigma_X\sigma_\Delta}) \tag{35}$$

.

Therefore, $P(Q_G^E = \overline{Q}) = P(Z > (b_G + \alpha + \phi)\overline{Q})$. For any normal random variable, $\frac{\partial P(Z > \overline{Z})}{\sigma_Z} > 0$. Therefore, $\frac{\partial P(Z > (b_G + \alpha + \phi)\overline{Q})}{\partial \sigma_X} > 0$, $\frac{\partial P(Z > (b_G + \alpha + \phi)\overline{Q})}{\partial \sigma_\Delta} > 0$ and $\frac{\partial P(Z > (b_G + \alpha + \phi)\overline{Q})}{\partial \rho} < 0$. From the previous discussions we know that $\mathbb{E}(\text{Spread}|Q^E < \overline{Q}) < \mathbb{E}(\text{Spread}|Q^E = \overline{Q})$ and $\mathbb{E}(\pi|Q^E < \overline{Q}) < \mathbb{E}(\pi|Q^E = \overline{Q})$, therefore $\mathbb{E}(\text{Spread}_{t+1})$ and $\mathbb{E}(\pi_{t+1})$ have a positive relation with $P(Q_{t+1}^E = \overline{Q})$. Hence, $\frac{\partial \mathbb{E}(\pi_{t+1})}{\partial \sigma_X} > 0$, $\frac{\partial \mathbb{E}(\pi_{t+1})}{\partial \sigma_\Delta} > 0$ and $\frac{\partial \mathbb{E}(\pi_{t+1})}{\partial \rho} < 0$. The same argument applies for spreads. Q.E.D

References

Aid, Rene, Luciano Campi, Adrien Nguyen Huu, and Nizar Touzi, 2009, A Structural Risk-Neutral Model Of Electricity Prices, *International Journal of Theoretical and Applied Finance* 12, 925–947.

Bessembinder, Hendrik, and Michael L. Lemmon, 2002, Equilibrium Pricing and Optimal Hedging in Electricity Forward Markets, *The Journal of Finance* 57, 1347–1382.

Blair, Benjamin F., and Jon P. Rezek, 2007, The Effects of Hurricane Katrina on Price Pass-through for Gulf Coast Gasoline, *Economic Letters* 98, 229–234.

Borenstein, S, and A Shepard, 2002, Sticky Prices, Inventories, and Market Power in Wholesale Gasoline Markets, *Rand Journal of Economics* 33, 116–139.

Carlson, M, Z Khokher, and S Titman, 2007, Equilibrium Exhaustible Resource Price Dynamics, *Journal of Finance* 62, 1663–1703.

Carmona, Rene, and Valdo Durrleman, 2003, Pricing and Hedging Spread Options, *Society for Industrial and Applied Mathematics* 45, 627–685.

Chesnes, Matthew, 2009, Capacity and Utilization Choice in the US Oil Refining Industry, Working Paper.

Considine, T. J., 2001, Markup pricing in petroleum refining: A multiproduct framework, *International Journal of Industrial Organization* 19, 1499–1526.

Dempstera, M.A.H., Elena Medovab, and Ke Tang, 2008, Long term spread option valuation and hedging, *Journal of Banking & Finance* 32, 2530–2540.

Du, Xiaodong, and Dermot J. Hayes, 2008, The Impact of Ethanol Production on U.S. and Regional Gasoline Prices and on the Profitability of the U.S. Oil Refinery Industry, *Working Paper*.

Duan, Jin-Chuan, and Stanley R. Pliskab, 2004, Option Valuation with Co-integrated Asset Prices, *Journal of Economic Dynamics and Control* 28, 727–754.

Ghoddusi, Hamed, 2010, Dynamic Investment in Extraction Capacity of Exhaustible Resources, *International Journal of Theoretical and Applied Finance* 57, 359–373.

Gilbert, Richard J., and Justine S. Hastings, 2005, Market Power, Vertical Integration and the Wholesale Price of Gasoline, *The Journal of Industrial Economics* pp. 469–492.

Jin, Yanbo, and Philippe Jorion, 2006, Firm Value and Hedging: Evidence from U.S. Oil and Gas Producers, *The Journal of Finance* 61, 893–919.

Kilian, L, 2009, Why Does Gasoline Cost so Much? A Joint Model of the Global Crude Oil Market and the U.S. Retail Gasoline Market, Working Paper.

Kogan, Leonid, Dmitry Livdan, and Yaron Amir, 2009, Oil Futures Prices in a Production Economy with Investment Constraints, *International Journal of Theoretical and Applied Finance* 64, 1345–1375.

Manoliu, Mihaela, and Stathis Tompaidis, 2002, Energy Futures Prices: Term Structure Models with Kalman Filter Estimation, *Applied Mathematical Finance* 9, 21–43.

Oladunjoye, Olusegun, 2008, Market structure and price adjustment in the U.S. wholesale gasoline markets, *Energy Economics* 30, 937–961.

Radchenko, S, 2004, Oil price volatility and the asymmetric response of gasoline prices to oil price increases and decreases, Working Paper.

Rajgopal, Shivaram, 1999, Early Evidence on the Informativeness of the SEC's Market Risk Disclosures: The Case of Commodity Price Risk Exposure of Oil and Gas Producers, *The Accounting Review* 74, 251–280.

Routledge, Bryan R., Chester S. Spatt, and Duane J. Seppi, 2001, The "Spark Spread:" An Equilibrium Model of Cross-Commodity Price Relationships in Electricity, *Tepper School of Business.* 28, 86–89.

Sadorsky, Perry, 2001, Risk Factors in Stock Returns of Canadian Oil and Gas Companies, *Energy Economics* 23, 17–28.

Schwartz, Eduardo, and James E. Smith, 2000, Short-Term Variations and Long-Term Dynamics in Commodity Prices, *Management Science* 46, 893–911.

Vuolteenaho, Tuomo, 2006, What Drives Firm-Level Stock Returns?, *The Journal of Finance* 57, 223–264.