3rd International Workshop on Empirical Methods in Energy Economics (EMEE2010)

Surrey Energy Economics Centre (SEEC)
University of Surrey, UK
24th – 25th June 2010

NOTE:

The following Abstract and/or Paper is *Work in Progress* for presentation and discussion at the EMEE2010 workshop. It therefore must not be referred to without the consent of the author(s).

Sponsored by:

The Transmission of Risk in the Energy Supply Chain

Hamed Ghoddusi, Sheridan Titman, Stathis Tompaidis *

Preliminary — Please Do Not Quote

*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.

Introduction

In their most abstract formulation, supply chains consist of three components: an upstream primary input; a downstream output product; and a capital asset, the employment of which turns the input into the output. One can give numerous examples: crude oil is transformed into heating oil and gasoline at a refinery; bauxite ore is transformed into aluminum in aluminum smelters; jet fuel is transformed into air travel using an airplane. In each case, characteristics of the supply of the input, the demand of the output, and the physical properties of the capital asset determine the behavior of the spread in prices between the input and output. In this paper we offer a model for this spread, based on realistic assumptions on the three components of supply chains.

Much of the intuition of our model is captured in an example discussed in the chapter XI of *Politics*, by Aristotle. Aristotle describes a transaction that proved particularly gainful for one of the sages of antiquity, Thales of Miletus. It is told that Thales correctly forecasted a bountiful year for olive production in ancient Greece. Based on this prediction, months ahead of the olive harvest, Thales made a down payment to secure the use of all the olive presses. When the large harvest materialized, with a large supply of olives, and a limited number of olive presses, Thales was able to rent out the olive presses for a large profit. Aristotle argues that Thales was able to make a profit by creating a monopoly. Our model offers an alternative explanation: assuming that the demand curve for olives and olive oil remain relatively unaffected by the supply of olives, a year of extraordinary large supply of olives would have several effects. Since olives are bountiful, the price of olive would drop; given the cheaper price of input, olive oil prices would also drop. However, once olive presses operate at capacity, the amount of olive oil available can no longer increase, and further drops in the price of the input are not reflected in the price of olive oil. At that point, the spread between olive prices and olive oil prices starts increasing as olive prices drop further. Widening spreads lead to larger profits for owners of olive presses, even if no one enjoys a monopoly.

The model we develop in this paper makes assumptions on the three components of the supply chain: we assume an exogenous process for the price of the upstream input, which corresponds to the equilibrium between the input's supply and demand; we assume an exogenous demand curve for the downstream output; and we assume constraints in the operation of the capital asset that transforms the input into the output. Our model allows us to make predictions on the dynamics of price of the output, including the spread between the prices of the input and the output, the volatility of the price of the output conditional on the price of the input, the correlation between the two prices, and derive a process for the value of the capital asset, and its volatility and correlations with the input and output prices. Assuming a link between the demand for the downstream product and the market return, our model can also produce conditional predictions for the beta coefficient between the returns of the output and the capital asset and the market portfolio.

Our model leads to several predictions: a) in a perfectly competitive market with no constraints in the operation of the capital asset, spreads converge to a fixed level, while when the operation of the capital asset is constrained — either through capacity limits, or production adjustment costs — spreads deviate from this level; b) when demand is high enough, spreads under both a monopoly and a competitive market converge; c) very low input prices, and stable output demand increases spreads. In this case the capital asset operates almost at full capacity; d) correlation between input and output prices depends on the level of output demand and input price. When output demand is low and input price is high, the price of the input and the price of the output are highly correlated. When output demand is high or input price is very low, the price of input and the price of the output diverge and the correlation drops significantly; e) the spread increases when the volatility of output demand increases. This happens because of convexity in the spread; f) correlation between input and output prices is higher in competitive markets than in monopolistic markets, and; g) output demand shocks result in bigger output price changes in a competitive market than in a monopolistic market.

In addition to developing an equilibrium model, we provide an empirical study of the model in the case where the input is crude oil, the output is gasoline, and the capital asset a refinery. Using a Kalman filter, we estimate the dynamics of gasoline demand under both the physical and risk-neutral measures using a time series of observed prices of gasoline and crude oil and by assuming that the gasoline market is competitive. We also provide comparative statics that explore how other exogenous parameters, such as the demand for gasoline, the supply of oil and the production function, affect the endogenous variances and correlations. Since we have daily spot price data from 1984, we also examine differences in the observed behavior of oil and gasoline prices as well as the returns of refineries in different periods. In the early part of our sample the market for refined products was very competitive. Starting in the late 1990s there was substantial consolidation of refineries, and near the end of our sample period the market was much less competitive. The sample periods also differ in the extent to which refineries were operating at capacity. In most of our sample period refineries were operating substantially below capacity. However, in the 2005-2007 period, production was much closer to capacity levels.

Our results illustrate the challenges associated with hedging input prices. Since in our model there are two state variables, the volatility of profits may not be meaningfully reduced by simply hedging input prices. Indeed, we show that small changes in parameters can lead to dramatic changes in the hedge ratio. It is possible that, due to the non-linearity of the relation between the downstream and upstream prices, non-linear hedging instruments, for example options, might provide a better hedge.

I. Literature Review and Empirical Evidence

A key issue in gasoline price is the asymmetry between shocks to the price of crude oil and the wholesale and retail price of gasoline. The asymmetry suggests that a positive shock to crude transmits much quicker to gasoline prices than a negative shocks.

Borenstein and Shepard (2000) provides explanations for why gasoline prices do not immediately react to the changes of crude price. Their model includes production adjustment costs and storage. There are costs and benefits associated with storage. Storage is costly on the one hand but decreases scheduling and distribution costs on the other hand. They use responses of gasoline futures price to the innovations of crude futures with the same maturity. If the lag between gasoline and crude prices is due to the slow adjustment of production, then near-to-maturity contracts of gasoline should react only partially while the long maturity contacts will immediately react with a full adjustment. The results show that 1 cent increase in crude price eventually increase gasoline price for 1.14 cents but this effect is 0.16 cents small for near-month future contracts.

Radchenko (2004) examines the hypothesis that the volatility of crude price affects the degree of asymmetric response of gasoline prices to oil price shocks. His findings are based on a VAR model that confirms a negative relationship between the volatility and the degree of price response asymmetry. Radchenko (2004) suggests that these findings may support search with Bayesian updating and oligopoly coordination theories in gasoline market.

There is a literature on the volatility spillover from input to output markets (vertical spillover). Buguk, Hudson, and Hanson (2003) build an EGARCH model of catfish price volatility by incorporating the lagged error terms of price equation for the production inputs (corn and soy). Blair and Rezek (2007) evaluate the effect of Katrina on the cost pass through of crude on gasoline. Using an ECM model, they show that in the period before, and the period long after the Katrina, every unit of increase in the price of crude oil has almost the same magnitude on gasoline prices. On the other hand, right after Katrina, a 10 cent increase in crude price, drives up the price of gasoline by 63 cents.

Kilian (2009) builds a structural VAR model to study the joint behavior of global crude price, US gasoline price, world production of crude oil, US gasoline consumption and the index of global economic activities. These five variables are used to separate the effects of five different shocks: 1)crude oil supply shocks 2) aggregate demand shocks 3) oil-market specific demand shocks 4) shocks to the supply of gasoline in the U.S. and 5) shocks to the U.S. demand for gasoline. His results suggest that: a) A negative shock to gasoline supply capacity will cause a persistent increase in gasoline price and a temporary decrease in the imported crude oil price; b) A positive shock to U.S gasoline demand takes two months to push gasoline prices up. It, however, has no significant effect on the price of crude oil. This finding suggests that inventories absorb demand shocks for a while until the producers realize the permanent shock. On the other hand, it does not affect global oil prices because of capacity constraints. Higher demand does not increase the production of gasoline much and therefore does not affect the demand for input (crude).

Dempster, Medova, and Tang (2008) propose a two-factor model for spread between two co-integrated assets. They directly model spread through a mean-reverting process with a stochastic mean-reverting long-run mean. They use state-state representation and Kalman filtering to calibrate the model and then apply it for valuation of futures and options on the spread between heating oil and Western Texas Intermediate (WTI) and also between WTI and Brent.

Kuper and Poghosyan (2008) use a Threshold Vector Error Correction Model (TVECM) to model the relationship between crude oil and gasoline prices. Their model shows that the co-integration relationship between crude and gasoline has a structural break in February 1999. The paper shows that before February 1999 the cointegration relationship between crude and gasoline was linear while after this period the relationship becomes non-linear.

Chesnes (2009) builds and estimates a dynamic model of refinery operation, maintenance, and investment. In his model, there is a trade-off between producing with high utilization rate and in longer time and bearing the risk of break-down in future periods. He uses a quadratic cost function to capture the effect of capacity utilization on production costs. The demand, break-down probability and cost functions are estimated using instrumental variables and GMM methods.

Considine (1997) and Considine (2001) provide insights on the modeling of sales, inventory and investment costs of the refinery sector. The model used in both papers assumes a quadratic adjustment cost for change in production and inventory level. The policy variables are the optimal production, sales and investment rate at each period. The first order conditions for the optimal inventory policy imply that the expected convenience yield minus interest and storage costs should be equal to the expected changes in production costs and adjustment costs.

A. Empirical Evidence

We discuss some empirical evidence to motivate our modeling of the crude oil supply and gasoline demand. Figure 1 shows the weekly price of gasoline and crude oil from January 2000 to September 2009.

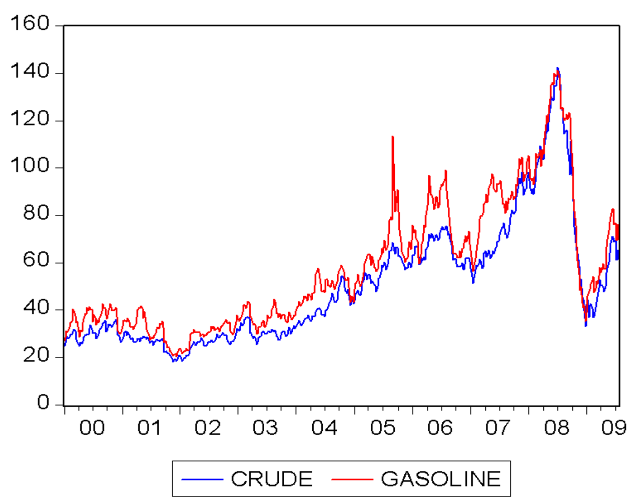

Figure 1. Weekly price of crude and gasoline

To see that the volatility of crude oil and gasoline move very closely together, Figure 2 depicts the GARCH series for crude oil and gasoline.

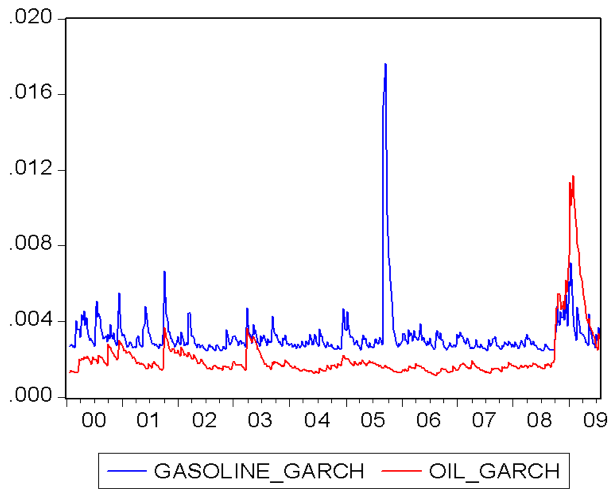

Figure 2. Garch Series for Crude and Gasoline

Variable	Coefficient	t-value
Constant	0.000343	0.2
CR	0.776	18.9
CR(-1)	0.114	2.8

Table 1
Correlation of Gasoline and Crude Returns

We regress weekly gasoline returns on the weekly returns of crude oil (denoted by CR) including one lag of crude return. Table 1 shows the result of the regression which suggest that there is a high correlation of instantaneous changes in crude and gasoline prices.

B. Equilibrium Relationship

To check for the existence of an equilibrium relationship between crude oil and gasoline, we estimate a vector error correction model (VEC) using the Johansen procedure and look into the cointegration relationship between gasoline and crude prices. If the cointegration relationship

Linear Form	Logarithmic Form
$P_G = 3.92 + 1.05P_C$	$p_G = 0.35 + 0.94p_C$

Table 2
Cointegration Relation between Gasoline and Crude

between gasoline and crude is defined as $Z_t = P_{G,t} - \alpha P_{C,t} - d$ (in level or log), the VEC equations will be:

$$\Delta P_{G,t} = \alpha_G + \beta_G Z_{t-1} + \sum_{i=1}^{J} a_{G,i} \Delta P_{G,t-i} + \sum_{i=1}^{J} b_{G,i} \Delta P_{G,t-i}$$
$$\Delta P_{C,t} = \alpha_C + \beta_C Z_{t-1} + \sum_{i=1}^{J} a_{C,i} \Delta P_{G,t-i} + \sum_{i=1}^{J} b_{C,i} \Delta P_{G,t-i}$$

Using two separate functional forms for levels and logarithms of levels, one gets the values reported in Table 2 for the cointegration relationship from the VEC procedure

We interpret the residuals from the cointegration relationship as proxies for latent demand process where one of the major goal of theoretical modeling and estimation procedure in the later sections is to have a reasonable estimation about it.

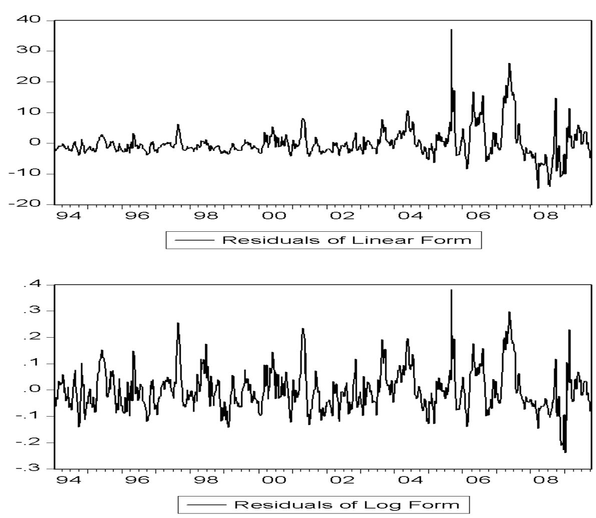

Figure 3. VEC Residuals

C. Refining Capacity and Crack Spreads

The term ”crack spread” refers to the difference between the price of crude oil and that of certain derivatives of crude oil (mostly gasoline and heating oil). There are synthetic contracts traded in NYMEX and other commodity exchanges which offer a 3-2-1 ratio, meaning that the value of the contract is the difference between the value of three units of crude oil and the sum of two units of gasoline and one unit of heating oil.[1]

Figure 4 shows the joint graph of the crack ratio; i.e., the ratio of gasoline to crude prices, and the percentage of active capacity utilization of U.S refineries. Regressing crack spreads on capacity utilization we find a significant positive coefficient with $R^2 = 26\%$. This result supports a basic economic model where the production level drops when the price of the input increases.

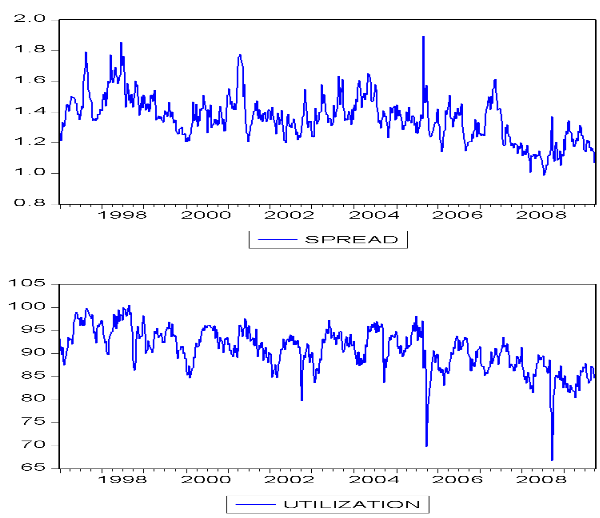

Figure 4. Spread and Capacity Utilization

Crack spreads follow a seasonal pattern due to the seasonality in gasoline demand. We calculate the average monthly crack spread from 1986-2010. Table 3 shows the normalized monthly spread factor.

[1]Although such a contract can be easily replicated by a portfolio of contracts on these three products, market participants prefer it because of its lower margin account requirements.

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
0.70	0.75	0.88	1.08	1.31	1.51	1.33	1.15	1.11	1.14	1.21	0.91

Table 3
Seasonality Factors

D. Principal Component Analysis

We run principal component analysis (PCA) on a panel of de-seasonalized crack spread forwards from 2005 to 2009. The forward values are extracted from the difference between crude and gasoline forward prices for the range of 1 to 12 months. The cumulative eigenvalues of PCA are presented in Figure 5. 90% of total variance can be explained by the first three principal components.

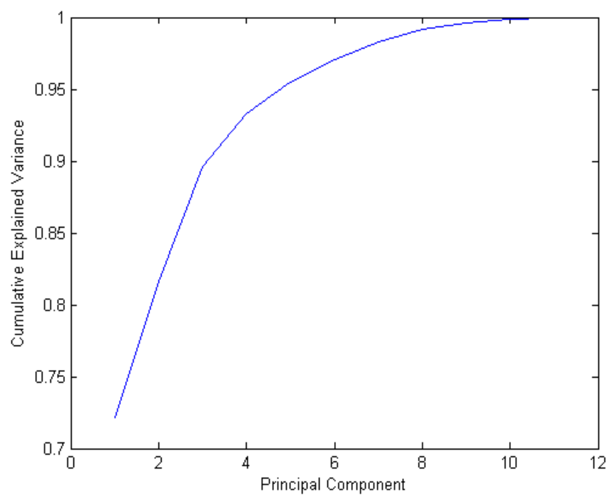

Figure 5. Cumulative Explained Variance

The weights associated with each contract (1:12) are presented in figure 6. The weights suggest that, the first principal component explains 72% of total variance and is a level variable. Two other variables are related to the slope of the short and long maturity contracts.

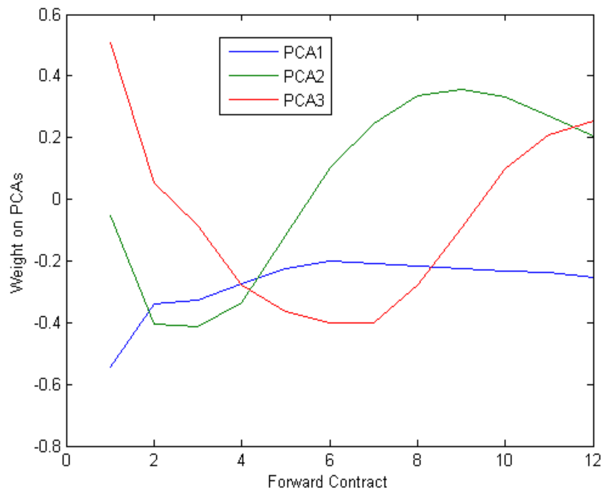

Figure 6. Weights on Each Forward Contract

E. Risk Premium

It is interesting to know the sign of the risk premium in the gasoline market. On the one hand, if buyers are more risk-averse, they are willing to pay a premium to get a pre-determined price of gasoline. In this case futures prices should be higher than realized spot prices. On the other hand, if sellers are more risk-averse they pay a risk premium in order to sell their output without risk, leading to lower futures prices than the realized spot prices.

We use the existing 12-month futures contracts for RBOB (New York harbor) starting from Oct 2005 in order to estimate the risk premium. Risk premium is defined as the difference between realized spot price at the maturity of a forward contract and the average price of all futures contracts maturing on that day. Figure 7 shows the trend of the risk premium expressed as a percentage. At the beginning of the data, the risk premium is close to zero, then becomes positive during the pre-crisis era (the time of the commodity spot price boom), turns negative during the crisis period and eventually recovers back to the positive region as the crisis ends. The graph may suggest that in normal times the risk premium is positive.

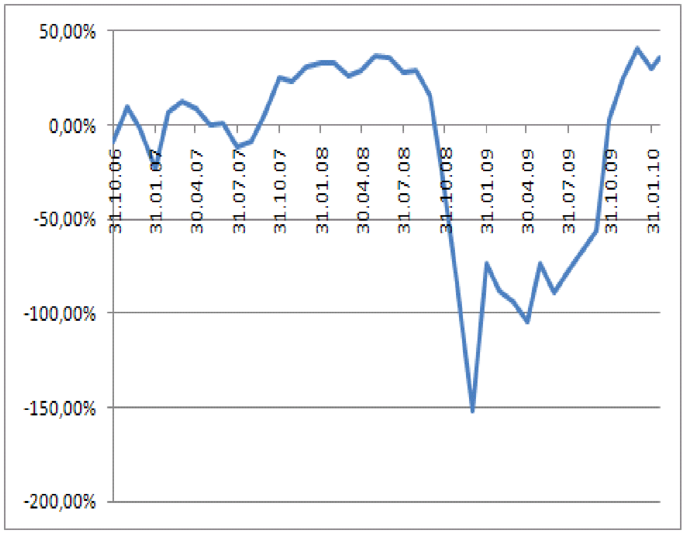

Figure 7. Risk Premium in Gasoline Futures Contracts

II. Model

We develop an equilibrium model for the relationship between the price of the input and the output and the operational characteristics of the capital asset that transforms the input into the output. The model is built on assumptions on each of these three components. Throughout we consider the case of crude oil as input, gasoline and heating oil as output, and a refinery as the capital asset that transforms the input into the output.

Input: Crude Oil

There are two sources of demand for crude oil in the global economy: U.S refinery demand, consisting of gasoline, diesel and heating oil, and all non-US crude demand. In the short run both demand factors are mean-reverting. Although both factors can potentially contain some seasonal factors, the literature and our empirical tests do not show a significant seasonal element in the price of crude oil. As a result, both factors of demand are treated as non-seasonal processes.

Crude supply is subject to random shocks with mean-reverting dynamics. 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. The impact of OPEC trigger-type policies, where supplies are increased when prices cross a certain price level and decreased when prices drop below a different level, drive the mean reversion in crude oil prices.

Given this background, we abstract from modeling demand for oil and directly model crude oil prices, P_C, as an Ornstein-Uhlenbeck mean-reverting process. Under this process, crude oil prices are log-normally distributed with mean and variance $(\mu_{P_{C,t}}, \sigma_{P_{C,t}})$, which can be calculated given the current value of the process and its parameters.

Output: Gasoline

The demand curve for gasoline in the United States will be assumed to be a linear function

$$P_G(q) = X - bq$$

where X is a random factor depending on variables such as income and weather conditions, b is a constant elasticity parameter, and q is the quantity of gasoline supplied to the market. While a linear demand function is unrealistic for gasoline, since it suggests the existence of a price above which demand for gasoline is zero, we use it as an approximation of the true demand function which allows us to derive analytical solutions. We assume that the logarithm of the deseasonalized gasoline demand X follows an Ornstein-Uhlenbeck mean-reverting process, resulting in demand that is log-normally distributed with mean and variance $(\mu_{X,t}, \sigma_{X,t})$. The correlation between the logarithm of gasoline demand and the crude price is given by $\rho = \mathrm{Cov}(X, P_C)$.

Capital Asset: US Refinery Industry

We assume that all the gasoline consumed in the US is produced domestically and there are no gasoline imports or exports. The domestic market is supplied by a refinery industry which converts crude oil into refined products including gasoline, heating oil and jet fuel. In reality, refineries may use different forms of crude — e.g. heavy crude and light sweet —

depending on their design and technical specifications. For simplicity, we do not distinguish between different input types and treat crude oil as a homogeneous commodity. We focus on a horizon of approximately one year. In this horizon, the refining capacity is known and no capacity can be added. Although refineries can make adjustments in their gasoline to distillate mix, historical data show that the ratio of gasoline to distillate production is approximately equal to two. We abstract from the product-portfolio optimization problem of the refinery and simply assume an aggregate demand and product model.

Perfect Competition

We assume that there exist a large number of homogeneous price-taker refiners with a total industry capacity of $\overline{q}$. The industry can adjust the operational capacity costlessly but can not produce above maximum capacity $\overline{q}$. The capacity constraint $\overline{q}$ is given exogenously. The refining cost consists of the price of the crude oil input, other inputs — energy, chemicals, wages, etc — and operational and capital costs. As refineries move to production levels close to maximum capacity, inefficiencies, such as higher probability of future break-downs, arise leading to an increase in the marginal cost.

The refining marginal cost function for production levels, q, less than maximum capacity, is given by:

$$MC(q) = P_I + P_C(q) + \phi \frac{q}{\overline{q}} \tag{1}$$

where P_I is the cost of other inputs, $\overline{q}$ is the maximum capacity of refinery and ϕ the parameter for capacity related costs. We assume that the marginal cost for production levels above maximum capacity is infinite.

Equilibrium Gasoline Price

When production is less than capacity, the optimal capacity level solves the first order conditions for profit maximization and we have that marginal profil equals marginal cost, which implies that the optimal level of production, q^*, is given by

$$q^* = \min\left\{\overline{q}, \frac{X - P_I - P_C}{b + \frac{\phi}{\overline{q}}}\right\} \tag{2}$$

The price of gasoline is given by:

$$P_G = \begin{cases} \frac{\frac{2\phi}{\overline{q}}X + b(P_C + P_I)}{b + \frac{\phi}{\overline{q}}} & \text{if } q^* \leq \overline{q} \\ X - b\overline{q} & \text{otherwise} \end{cases}$$

III. Model Calibration and Estimation of Parameters

A. Objectives and Data

In a competitive market, there is not much to estimate regarding the spreads between input and output. Dynamic spread is generated when there are some deviations from perfect competition. As pointed out before, we do not know the true economic model of refinery industry. However, it was shown that if one assumes a major type of friction (market power, capacity constraints, capacity costs, adjustment costs) one arrives at a similar functional form for the relationship between gasoline and crude. In all of these models gasoline price is determined by some fuction of crude price and (unobserved) demand shock. A full specification of relationship between two prices requires assumptions regarding the functional form of demand process, too. With a linear demand, one can propose a general reduced form of $P_G = d + m_1 P_C + m_2 X$ where $\{d, m_1, m_2\}$ are unknown coefficients to be estimated. In a perfectly competitive market $m_2 = 0$.

We calibrate our simple model to gain some intuition over the magnitude and dynamics of the key variables. The input of calibration process consists of observable variables (gasoline price, crude price, gasoline production or $\{P_G, P_C, q\}$) and the goal is to estimate the values for structural parameters ($\{m_1, m_2, P_I\}$) as well as the dynamics of demand process ($\left\{\mu_X, \overline{X}, \sigma_X\right\}$) under both the physical and risk-neutral measures.

Our data set consists of weekly observations of crude and gasoline price and gasoline production from 1990/11/02 to 2010/04/23 (1017 observations in total), obtained from EIA website. For futures prices we use Bloomberg data for contracts traded in NYMEX between 2005/10/07 and 2009/09/25 (218 observations).

B. Transition, Measurement and Updating Equations

Using standard methods of state-space representation and leting $Z_t = P_G - m_1 P_C$ or $Z_t = P_G - m_1 q$, the transition and measurement equations of Kalman filter is defined as:

$$
\begin{aligned}
X_{t+1} &= c_X + HX_t + \varepsilon_1 \\
Z_t &= FX_t + d\varepsilon_2 \\
c_X &= \mu_X \overline{X}, H = (1 - \mu_X) \\
F &= m_2 \\
\mathbb{E}(\varepsilon_1) &= 0 \\
\mathbb{E}(\varepsilon_2) &= 0 \\
\sigma_T &= \mathbb{E}(\varepsilon_1^2) \\
\sigma_M &= \mathbb{E}(\varepsilon_2^2) \\
< \varepsilon_1, \varepsilon_2 > &= 0
\end{aligned}
\tag{3}
$$

The prediction equation and estimated covariance used for filtering is given by

$$\hat{X}_{t+1|t} = \mathbb{E}(X_{t+1}|Z_t) = c_X + H\hat{X}_t + A_{t+1}(Z_{t+1} - \hat{Z}_{t+1})$$

$$\hat{Z}_t = F(c_X + H\hat{X}_t) + d$$

$$A_t = R_t F_t Q_t^{-1}$$

$$R_t = H\Sigma_{t-1}H + \sigma_T$$

$$\Sigma_t = R_t - A_t Q_t A_t^{'}$$

$$Q_t = (F_t R_t F_t + \sigma_M)$$

To estimate the unkown parameters the likelihood function of the Kalman filter should be maximised. Since the errors are normally distributed, the log-likelihood function can be calculated using each period's estimated variance R_t and the error $e_t = Z_t - F\hat{X}_{t|t-1}$

$$L(\theta, P_0) = \sum_{t=2}^{t=T} -\frac{1}{2}\ln(|R_t|) - \frac{1}{2}e_t R_t^{-1} e_t$$

where P_0 is the prior regarding the mean and variance of X_0. The parameters can be found by minimizing $L(\theta, P_0)$ using numerical procedures in Matlab.

C. Estimation of Parameters Under Physical Measure

The unkown parameter set of our problem consists of $\theta = [m_1, m_2, d, \overline{X}, \mu_X, X_0]$. We run the Kalman filter for two different specifications

Price Model

The first specification is the reduced-form relation between gasoline and crude prices. To include the cost of other inputs P_I into the estimation, we assume that $P_I = \alpha P_C + d$. Part of the cost is proportional to crude price and is captured by the $(1+\alpha)P_C$ term. The remaining

m_1	m_2	d	μ	$\overline{X}$	X_0
1	3.55	-0.58	0.09	1.86	1.95

Table 4
Parameter Estimation for the Entire Period

part appears as a fixed term in the estimation equation. The economic model reduces to the following equation to be estimated:

$$P_G = m_1 P_C + m_2 X + d + \varepsilon_1, \varepsilon_1 = N(0, \sigma_M)$$

where P_G and P_C are observable, X is an unobservable state variable and $\{m_1, m_2, d, X_0, \mu_X, \overline{X}\}$ are parameters to be estimated.

The Kalman filter is first applied to the entire sample between 1990-2010 (1017 observations). The estimated parameters are reported in table 4. Graph 8 shows the filtered values of unobservable variable X. From the filtered values we estimate the variance of demand process, $\sigma_X = 1.7$.

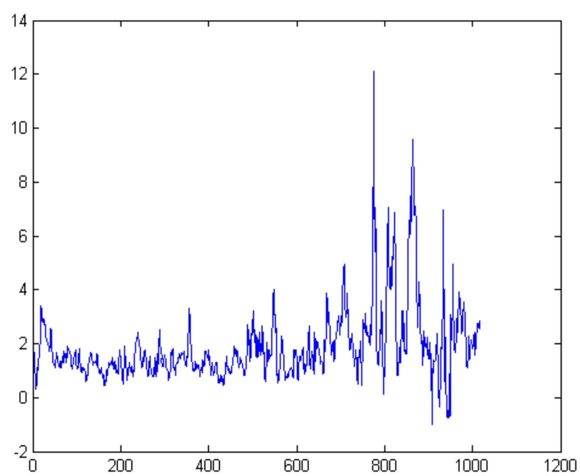

Figure 8. Estimated Values of Demand Process (X), 1990-2010

Year	m_1	m_2	d	μ	$\overline{X}$	X_0
1991	0.62	1.40	-0.040	0.053	9.69	9.50
1992	1.01	0.83	-0.64	0.19	5.51	3.96
1993	1.26	0.69	-3.20	0.085	2.19	1.77
1994	0.94	1.19	-0.88	0.28	4.54	3.47
1995	1.48	1.06	-1.70	0.076	-2.93	-3.14
1996	0.56	1.27	-3.10	0.13	13.49	10.11
1997	0.98	1.39	0.48	0.10	3.40	1.77
1998	1.00	0.73	0.38	0.053	3.27	4.02
1999	1.41	0.57	-21.39	-0.03	35.79	31.97
2000	1.1	2.36	0.00	0.26	1.29	-0.76
2001	1.06	2.33	-0.67	0.082	1.57	2.42
2002	0.92	1.57	-0.26	0.21	4.36	2.78
2003	0.93	2.21	-0.41	0.17	3.95	2.52
2004	1.13	2.37	0.45	0.022	-1.75	1.16
2005	1.79	8.95	1.75	0.48	-4.22	-3.43
2006	1.49	4.67	-0.12	0.15	-4.61	-4.78
2007	0.69	4.46	-1.66	0.09	9.23	4.84
2008	0.61	6.69	-2.80	-0.00	47.51	6.67
2009	0.85	3.69	-1.91	0.35	5.20	1.29

Table 5
Parameter Estimation Annually

The procodure is repeated for each year between 1990-2010. The estimated parameters for annual estimations are reported in table 5

Demand Function Model

The second specification is based on the equation for the demand function of gasoline given as $P_G = X - bq +_2, \varepsilon_1 = N(0, \sigma_M)$. The problem with this specification is that we do not observe the actual net supply of gasoline to the market at each week. What is reported in the data is the weekly production of refinery sector while the net supply is the sum of refinery production and net changes in gasoline inventories. A complete specification requires modeling of dynamic production-storage problem. Table shows estimated values of b in the equation $P_G = m_2 X + m_1 Q + d$ for entire sample (table 6).

m_1	m_2	d	μ	$\overline{X}$	X_0
9.01	4.79	-0.07	0.00	12.90	6.30

Table 6
Parameter Estimation for the Entire Period

D. Risk-Neutral Dynamics

We use the futures prices of crude and gasoline in order to estimate market's belief regarding future values of X under risk-risk measure. The observable variables at each time point are two vectors of 12 futures contracts with maturities between 1 and 12 months for gasoline and crude. This approach allows us to jointly estimate the dynamics of X under both physical and risk-neutral measures and extract what we call the *term-structure of demand process*.

Investigation of the shape of crude and gasoline futures suggests — see figure 9 — that unlike crude, gasoline futures have an strong seasonal element. Therefore, we define $Q = [Q(T)], T \in \{1, ..., 12\}$ as a vector of montly factors to capture the seasonality effect of each maturity date T.

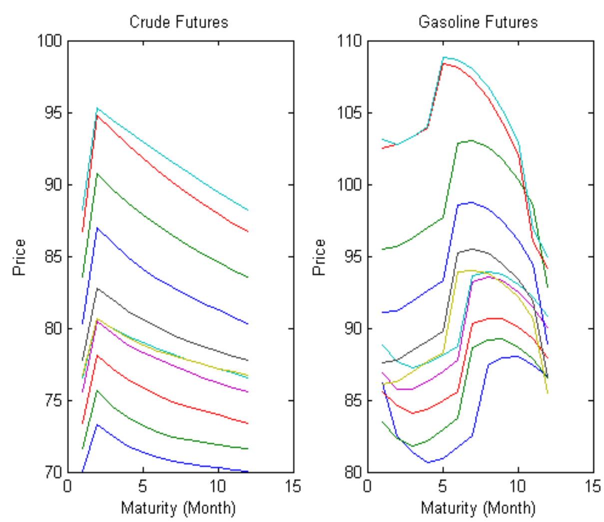

Figure 9. Estimated Values of Demand Process (X), 1990-2010

As before, we assume a single short-term mean-reverting factor for gasoline demand. Following other papers; e.g., Schwartz and Smith (2000) and Manoliu and Tompaidis (2000), we use the standard formulation as follows. Denote by $X(t)$ the value of stochastic demand parameter at time t. The dynamics of this process under Q-measure is given by

$$dX = \mu_Q(\overline{X}_Q - X)dt + \sigma_Q dW$$

Taking the expectation and multiplying by the relevant seasonality factor gives the expected value of X at any future maturity time T.

$$E(X(t,T)) = Q(T)(X(t)e^{-\mu_Q(T-t)} + \overline{X}_Q(1-e^{-\mu_Q(T-t)}))$$

If we rely on the linear structure for the relationship between gasoline and crude prices, then

$$E^Q_{t,T}(P_G) = m_1 E^Q_{t,T}(P_C) + m_2 E^Q_{t,T}(X)$$

For any given set of initial values and parameters $\{Q(T), X(t), \mu_Q, \overline{X}_Q\}$ and observed crude prices, F_{CL}, expected gasoline price $E^Q_{t,T}(P_G)$ and the estimation error $e_{t,T} = F_{Gas}(t,T) - E^Q_{t,T}(P_G)$ can be calculated. The Kalman filter is used to find the parameters of P and Q dynamics to maximize the likelihood function which depends on all $e_{t,T}$.

m_1	m_2	μ_Q	$\overline{X}_Q$	μ_X	$\overline{X}$	X_1
1.0378	3.3972	-0.0000	58.3293	0.0535	3.0249	1.6231

Table 7
Joint Physical and Risk-neutral Parameter Estimation

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1.00	0.87	1.20	1.07	0.90	0.79	0.81	0.89	1.13	0.87	0.86	0.88

Table 8
Monthly Factors

To summarize, the state-space representation of problem is the following

$$Y_t = M_1 d_t + M_1 X + M_2 Z$$
$$Y_t = [F_{Gas}(t,T)]$$
$$Z = [F_{Crude}(t,T)]$$
$$d_t = [Q(T)(\overline{X}_Q(1 - e^{-\mu_Q(T-t)}))]$$
$$X = [Q(T)X_t e^{-\mu_Q(T-t)}]$$
$$X_{t+1} = X_t e^{-\mu_X} + (1 - e^{-\mu_X})\overline{X}$$

The estimation results are summarized in Tables 7 and 8. $\mu_Q = 0$ and $Q(T) > 0, \forall T \in \{1,..,12\}$ suggest that the expectation regarding the future values of X under risk-neutral measure is captured by the seasonality factors. More formally,

$$\mu_Q = 0 \Rightarrow dX = \sigma_Q dZ \Rightarrow E(X(t,T)) = Q(T)X(t)d$$

IV. Model Extensions

In the previous sections, we presented a basic version of the model which assumed an exogenous oil price, fixed capacity, and perfect competition. In this section we relax these assumptions and study the effects of introducing competition in refining crude oil and frictions on the behavior of price and crack spreads.

A. Endogenous Crude Price

Crude Market

Let x_{US}, and q_{NUS} be the log of de-seasonalized demand for US refinery products and the net global crude supply. Their mean-reverting dynamics are given by:

$$
\begin{aligned}
dx_G &= \mu_G(\bar{x}_{US} - x_{US})dt + \sigma_{US} dW_{US} \\
dq_{NUS} &= \mu_{NUS}(\bar{q}_{NUS} - q_{NUS})dt + \sigma_{NUS} dW_{NUS} \\
< W_{NUS}, W_{US} > &= \rho
\end{aligned}
$$

The dynamics of demand and supply are given by the product of two deterministic monthly seasonal factors, f_{US} and f_{NUS}, and the stochastic mean-reverting factors.

$$
\begin{aligned}
X_{US} &= exp\{x_{US} f_{US}\} \\
Q_{NUS} &= exp\{x_{NUS} f_{NUS}\} \\
f_{US} &= (K_{US} + k_1 sin(\frac{\phi_{US}(t)\pi}{12})) \\
f_{NUS} &= (K_{NUS} + k_2 sin(\frac{\phi_{NUS}(t)\pi}{12}))
\end{aligned}
$$

The Crude Price

While, in reality, the global oil market consists of OPEC, large producers such as US and Russia and a competitive fringe of small producers, we assume that the market is competitive with an increasing marginal cost curve. The intuition behind the positive slope for marginal-cost curve is based on the heterogeneity in oil sources: when demand/supply ratio is low, high-cost small wells shut down and the price drops to the level of scarcity rent; when demand is strong enough, low cost sources are working at full capacity and some high-cost sources (including other sources such as oil shales) join the supply side and the price is determined by the cost of the marginal producer.

$$P_C = \alpha(Q_{NUS} + Q_{US})$$

B. Competitive Refinery with Endogenous Crude Price

As before, there exist a large number of homogeneous price-taker refiners with a total industry capacity of $\overline{q}$. The industry can adjust the operational capacity costlessly but can not produce above maximum capacity $\overline{q}$. The first order conditions give:

$$q^* = \min\left\{\overline{q}, \frac{X - P_I - \alpha Q_{NUS}}{b + 2\alpha + \frac{2\phi}{\overline{q}}}\right\}$$

The price of gasoline is given by:

$$P_G = X - bq^* = \begin{cases} \frac{(\frac{2\phi}{\overline{q}}+2\alpha)X+b(\alpha Q_{NUS}+P_I)}{b+2\alpha+\frac{2\phi}{\overline{q}}} & \text{if } q^* \leq \overline{q} \\ X - b\overline{q} & \text{Otherwise} \end{cases}$$

Crude oil price is determined via the feedback from the US gasoline market

$$P_C = \begin{cases} \alpha(\frac{X-P_I+(\alpha+b+\frac{2\phi}{\overline{q}})Q_{NUS}}{b+2\alpha+\frac{2\phi}{\overline{q}}}) & \text{if } q^*_{US} \leq \overline{Q}_{US} \\ \alpha(Q_{NUS} + \overline{Q}_{US}) & \text{otherwise} \end{cases}$$

Crack spreads are given by

$$CS_{\text{Competitive}} = \begin{cases} \frac{(\frac{2\phi}{\overline{q}}+\alpha)X+(b+\alpha)P_I-(\alpha^2+\frac{2\alpha\phi}{\overline{q}})Q_{NUS}}{b+2\alpha+\frac{2\phi}{\overline{q}}} & \text{if } q^* \leq \overline{q} \\ X-(b+\alpha)\overline{Q}_{US}-\alpha Q_{NUS} & \text{Otherwise} \end{cases}$$

C. Monopolist Refiner with Endogenous Crude Price

A monopolist refinery sector has market power and considers the effect of its production decisions on the equilibrium prices in the global oil market (input) as well as US refinery products (output). The monopolist choses q^* to maximize its profit according to the following equation:

$$q^* = \frac{X-\alpha Q_{NUS}-P_I}{2b+2\alpha+\frac{2\phi}{\overline{q}}}$$
$$Q_{US} = \min\left\{\overline{q}, \frac{X-\alpha Q_{NUS}-P_I}{2b+2\alpha+\frac{2\phi}{\overline{q}}}\right\}$$

The equation suggests that there is a singularity in the supply function of the refinery sector. When the domestic demand is too strong compared to net global oil demand-supply, the industry works close or at the full capacity and the production rate is independent of the crude oil market.

The price of gasoline is given by:

$$P_G = \begin{cases} \frac{(b+2\alpha+\frac{2\phi}{\overline{q}})X+\alpha b Q_{NUS}+bP_I}{2b+2\alpha+\frac{2\phi}{\overline{q}}} & \text{if } q^* \leq \overline{q} \\ X-b\overline{q} & \text{Otherwise} \end{cases} \tag{4}$$

According to Equation 4, and in line with empirical evidence, when US demand is high, refineries work at maximum capacity and US demand shocks do not influence crude oil price. When refineries operate at capacity the co-movement between US gasoline prices and global crude prices is lower than when refineries operate below capacity.

The crude price is determined by

$$P_C = \begin{cases} \alpha(\frac{X - P_I + (\alpha + 2b + \frac{2\phi}{\overline{q}})Q_{NUS}}{2b + 2\alpha + \frac{2\phi}{\overline{q}}}) & \text{if } q^*_{US} \leq \overline{Q}_{US} \\ \alpha(Q_{NUS} + \overline{Q}_{US}) & \text{Otherwise} \end{cases}$$

Crack spreads are given by

$$CS_{\text{Monopoly}} = \begin{cases} \frac{(b + \alpha + \frac{2\phi}{\overline{q}})X + (b+\alpha)P_I - (\alpha^2 + \alpha b + \frac{2\alpha\phi}{\overline{q}})Q_{NUS}}{2b + 2\alpha + \frac{2\phi}{\overline{q}}} & \text{if } q^* \leq \overline{q} \\ X - (b+\alpha)\overline{Q}_{US} - \alpha Q_{NUS} & \text{Otherwise} \end{cases}$$

V. Comparative Statics

Proposition 1. If global crude oil price depends on US refinery demand, the difference between crack spread of monopoly and competitive market is ambiguous. If crude oil price is exogenous, crack spreads are always higher in monopoly market.

Proof 1. Computing the difference between crack spreads of monopoly and competitive market:

$$CS_{\text{Monopoly}} - CS_{\text{Competitive}} = \frac{b(\alpha + b)[(X + P_I) - \alpha Q_{NUS}]}{(2b + 2\alpha + \frac{2\phi}{\overline{q}})(b + 2\alpha + \frac{2\phi}{\overline{q}})}$$

The numerator in the fraction depends on the realized values of X and Q_{NUS} and can be positive or negative. If crude oil price is exogenous, αQ_{NUS} disappears and the fraction is always positive.

Proposition 1. Higher demand or crude price volatility increases the expected spreads. The effect is stronger when the mean demand is higher

Proof 2. In a competitive market the price of output is equal to input price as long as the capacity constraint is not binding. When the capacity binds, the demand shocks become relevant and the difference between price output and input increases.

A time-varying volatility regime for gasoline demand and crude price is implemented. Figure 10 shows the effect of gasoline demand volatility on crack spreads. In the low volatility regime, the probability of hitting capacity constraint is low and therefore crack spread is smaller. While in high volatility regime, the price of crude and gasoline are more disconnected and crack spread increases. Observe that the crack spread under competitive market is systematically smaller than monopoly.

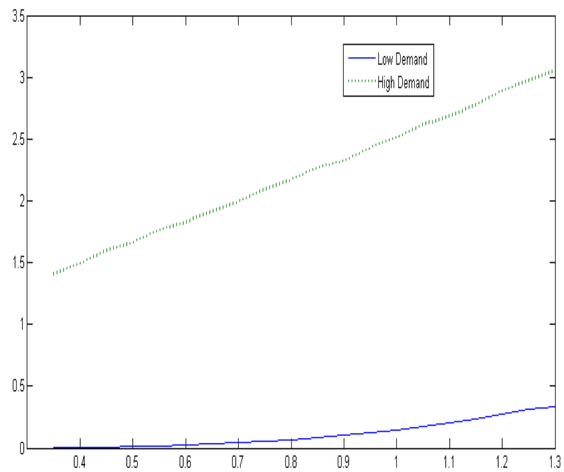

Figure 10. Effect of Demand Volatility on Crack Spreads

Proposition 1. With a fixed demand parameter, decreasing input price (crude in this case) will increase crack spreads after some point.

Intuitively, when crude price is low gasoline production increases and therefore the probability of meeting capacity constraints goes up. The model suggests that when crude price is very low, crack spreads increases quickly because gasoline price can not be lower than a certain amount dictated by capacity constraints. This result can be empirically tests in a regression of $\frac{\text{Gasoline}}{\text{Crude}} = a + b\text{Crude}$

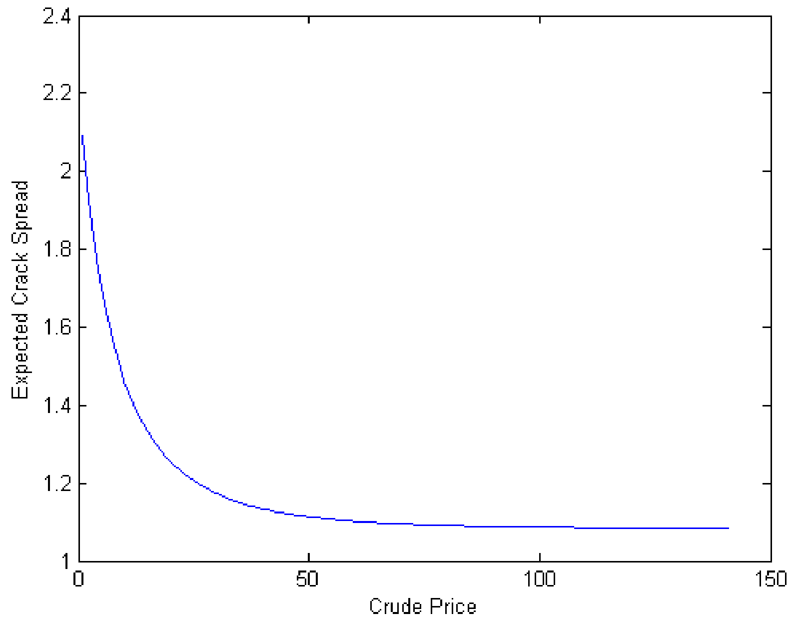

Figure 11. Crack Spreads vs Crude Price

VI. Probability Distribution of Gasoline Prices

It is of interest to know how the non-linearity in the supply function caused by capacity constraints determine the probability distribution of output prices. In this section, we introduce a computational method to calculate the conditional probability of gasoline prices given the first and second moments of crude price and demand parameter. We assume that there is also no correlation between crude oil price and the U.S gasoline demand. In the later sections we will relax these assumptions and solve a more complicated model where the crude price and gasoline demand are correlated and refineries bear a cost to produce at high capacity.

A. Assumptions for the Simple Case

The log of de-seasonalized gasoline demand X and crude price P_C both follow an Ornstein-Uhlenbeck mean-reverting process. Therefore, at each time point t in future their value are distributed according to a log-normal distribution with the mean and variance of $(\mu_{X,t}, \sigma_{X,t})$ and $(\mu_{P_{C,t}}, \sigma_{P_{C,t}})$ which can be calculated given the current value of the process and its parameters. The correlation between gasoline and crude price $\rho = Cov(X, P_C)$ is not necessarily zero but for simplicity will be assumed to be zero. The capacity constraint $\overline{q}$ is given exogenously. The market is competitive, demand is linear $P_G = X - bq$ and the optimal production rate is given by $q^* = \text{Min}\left\{\overline{q}, \frac{X-P_C-P_I}{b}\right\}$. The price is given by

$$P_G = \begin{cases} P_C + P_I & \text{if } q^* \leq \overline{q} \\ \\ X - b\overline{q} & \text{Otherwise} \end{cases} \tag{5}$$

B. Conditional Probability of Gasoline Prices

Initially notice that in a perfectly competitive industry with no frictions and unbounded capacity, the distribution of gasoline prices would be the same as of the crude prices plus a constant equal to the cost of other inputs (which is assumed to be non-random). Whereas, when capacity constraints are introduced the distribution of gasoline prices is not trivial anymore. Under these conditions, any given gasoline price $\overline{P}$ can be the result of production decisions in two different regions of production function. Figure 12 provides a graphical view of the event space for gasoline price. The 45-degree line represents the locus of values for crude oil prices and gasoline demand factor where the capacity constraint starts binding. In the region below the line (region 1), the optimal production rate is lower than capacity constraint and the market price is equal to marginal cost (which is the crude price). In the region 2 capacity constraint is binding and the price is given by $P_G = X - b\overline{q}$ equation. As mentioned before, any given gasoline price $\overline{P}$ can potentially be given by a set of points on both of these two regions. The probability of price realization can therefore be computed by calculating the probability of these two disjoint events and we can calculate the exact probability distribution function

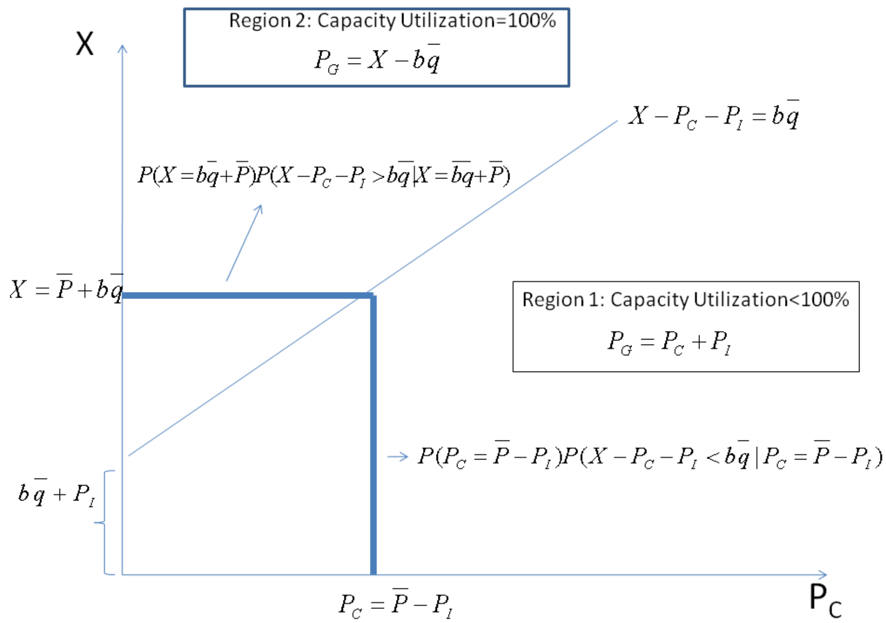

Figure 12. Event Space of Gasoline Prices

(PDF) of gasoline prices by calculating the probability mess of these two lines for each value of P_G.

Two vertical and horizontal lines in this graph show the set of values for crude price and gasoline demand which lead into a certain gasoline price $\overline{P}$. The density of gasoline price $\phi_G = P(P_G = \overline{P})$ can be decomposed into two elements: The first probability comes from production at interior region where price is equal to marginal cost. The second element of marginal density is generated by the binding region. Since the probability of being in the interior and binding regions (regions one and two) has a multivariate log-normal distribution depending on both P_C and X, one has to multiply the marginal distributions by conditional probability of being in that particular region. The optimal production rate is given by $\frac{X-P_C-P_I}{b}$; therefore given a value $\overline{P}_C$ for the crude oil and a fixed capacity level $\overline{q}$, the conditional probability of staying in the interior is $P(0 < X < \overline{P_C} + P_I + b\overline{q})$. Similarly, if the production takes place in the binding region, the price is given by $P_G = X - b\overline{q}$ and $X = b\overline{q} + \overline{P}$ generates gasoline price of $\overline{P}$ as long as $X - P_C - P_I > b\overline{q}$. Formally, we can write the PDF of gasoline price as:

$$\begin{aligned}
&\phi_G = P(P_G = \overline{P}) = \\
&P(P_C = \overline{P} - P_I)P(X - P_C - P_I < b\overline{q} | P_C = \overline{P} - P_I) + \\
&P(X - b\overline{q} = \overline{P})P(X - P_C - P_I > b\overline{q} | X = \overline{P} + b\overline{q}) = \\
&P(P_C = \overline{P} - P_I)P(X < b\overline{q} + \overline{P}) + P(X = \overline{P} + b\overline{q})P(P_C < \overline{P} - P_I)
\end{aligned} \tag{6}$$

Since both P_C and X follow a log-normal distribution, $P(P_C = \overline{P} - P_I)$ and $P(X = \overline{P} + b\overline{q})$ are easily given by the PDF of log-normal distribution (denoted by Pdf_X and Pdf_{P_C}). Furthermore, $P(X - P_C - P_I < b\overline{q} | P_C = \overline{P} - P_I) = P(X < b\overline{q} + \overline{P})$ and $P(X - P_C - P_I > b\overline{q} | X = \overline{P} - P_I + b\overline{q}) = P(P_C < \overline{P} - P_I)$ are determined from the formula for the cumulated density function (CDF) of log normal distribution which unfortunately does not have an analytical form and uses the error function.

$$\begin{aligned}
&P(P_G = \overline{P}) = \frac{1}{(\overline{P} - P_I)\sigma_C\sqrt{2\pi}} e^{-\frac{(ln(\overline{P} - P_I) - \overline{p}_C)^2}{2\sigma_C^2}} \phi\left(\frac{ln(\overline{P} + b\overline{q}) - \mu_X}{\sigma_X}\right) + \\
&\frac{1}{(b\overline{q} + \overline{P})\sigma_X\sqrt{2\pi}} e^{-\frac{(ln(b\overline{q} + \overline{P}) - \overline{X})^2}{2\sigma^2}} \phi\left(\frac{ln(\overline{P} - P_I) - \mu_{P_C}}{\sigma_{P_C}}\right))
\end{aligned} \tag{7}$$

Notice that the calculations of this section provide probabilities of having a fixed crack spread (region 1) or an increasing crack spread (region 2). In other words, these formulas enables us to calculate the conditional distribution of crack spreads given a certain gasoline price.

C. Results and Comparative Statics

Figure 13 depicts the probability distribution of gasoline prices for two different levels of capacity constraints. Comparing two graphs suggests that the effect of change in capacity

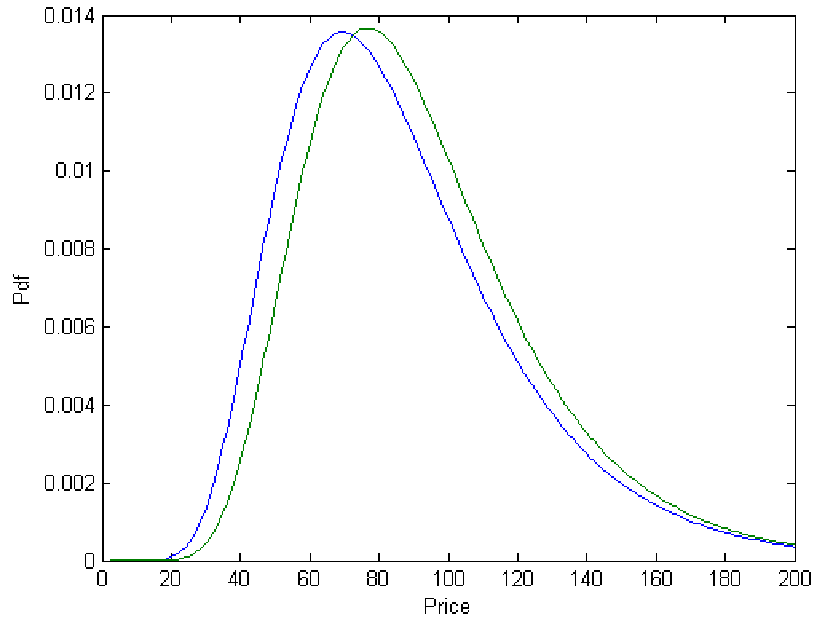

Figure 13. Probability Distribution Function of Gasoline Prices for Two Levels of Capacity

constraint on the marginal distribution of gasoline prices is not trivial. We can look at the effect of increasing capacity limit more formally:

$$\begin{aligned}
&P(P_G = \overline{P}|\overline{q}) = Pdf_{P_C}(\overline{P} - P_I)Cdf_X(b\overline{q} + \overline{P}) + Pdf_X(\overline{P} + b\overline{q})Cdf_{P_C}(\overline{P} - P_I) \\
&P(P_G = \overline{P}|\overline{q} + dq) = Pdf_{P_C}(\overline{P} - P_I)Cdf_X(b\overline{q} + bdq + \overline{P}) + \\
&Pdf_X(\overline{P} + b\overline{q} + bdq)Cdf_{P_C}(\overline{P} - P_I) \\
&\Rightarrow \frac{\partial \phi_G}{\partial \overline{q}} = Pdf_{P_C}(\overline{P} - P_I).Pdf_X + Cdf_{P_C}(\overline{P} - P_I)\frac{Pdf_X(K)}{\partial K}
\end{aligned} \tag{8}$$

The equations suggest that since the cumulative distribution is monotonically increasing, the first term is always positive: $\frac{\partial(Pdf_{P_C}(\overline{P}-P_I)Cdf_X(b\overline{q}+\overline{P}))}{\partial \overline{q}} = Pdf_{P_C}(\overline{P} - P_I).Pdf_X > 0$. However, $\frac{\partial(Pdf_X(\overline{P}+b\overline{q})Cdf_{P_C}(\overline{P})-P_I)}{\partial \overline{q}} <> 0$ because the PDF is not monotone and can increase or decrease. Therefore, adding to capacity constraints can increase or decrease the probability of certain value. If the PDF is a single peak one, then for values below its mode the term increases and for the terms after that it decreases.

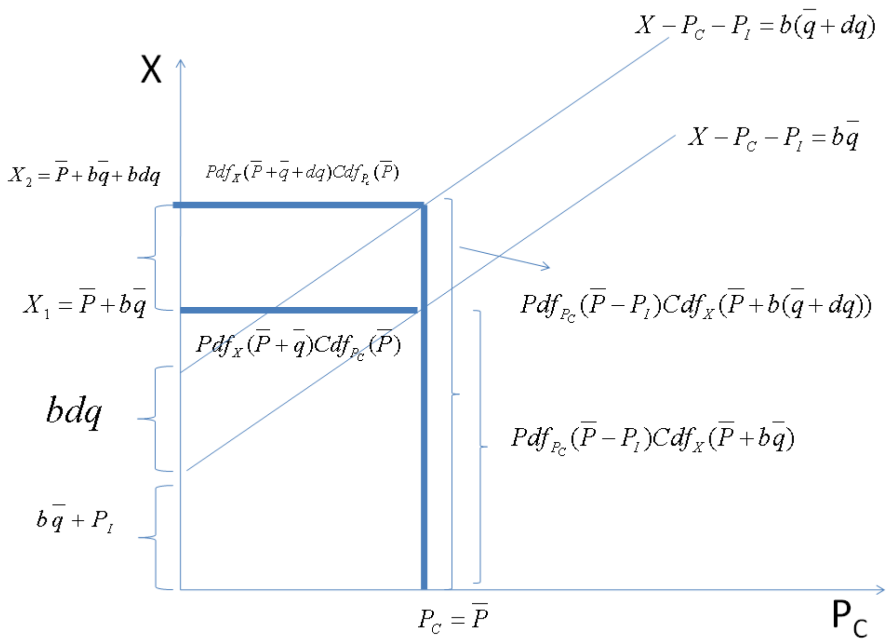

Figure 14. The Effect of Increasing Capacity

Now we can use the formula for ϕ_G to calculate the expected gasoline price. $\mathbb{E}(P_G) = \int_0^\infty P_G \phi_G(P_G) dP_G$. We discretize the state space and use a numerical procedure to estimate the expectation and variance of gasoline prices.

D. Correlated Model

We relax two key assumptions of the previous subsection and try to derive efficient approximations of gasoline price distribution under more realistic conditions.

D.1. Assumption

Crude oil price and gasoline price follow the same mean-reverting process as before. Unlike the previous section, they are correlated with each other with a coefficient ρ. Moreover, there are capacity related costs and market is not fully competitive. The quantity and price are given by

$$q^* = \text{Min}\left\{\overline{q}, \frac{X - P_I - P_C}{b + \frac{2\phi}{\overline{q}}}\right\}$$

$$P_G = \begin{cases} \frac{\frac{2\phi}{\overline{q}}X + b(P_C + bP_I)}{b + \frac{2\phi}{\overline{q}}} & \text{if } q^* \leq \overline{q} \\ \\ X - b\overline{q} & \text{Otherwise} \end{cases}$$

D.2. Probability Distribution of Gasoline Prices

The basic structure of the estimation procedure remains the same as before but some modifications are necessary. The condition $X - P_C \leq b\overline{q} + 2\phi + P_I$ ensures that the production takes place in the interior. The event associated with production in the interior and market price equal to $\overline{P}$ is defined as

$$E_1 = \left\{(X, P_C) : (X - P_C \leq b\overline{q} + 2\phi + P_I) \wedge (\frac{\frac{2\phi}{\overline{q}}X + b(P_I + P_C)}{b + \frac{2\phi}{\overline{q}}} = \overline{P})\right\} \Rightarrow$$

$$P(E1) = P((X - P_C \leq b\overline{q} + 2\phi + P_I) | \frac{\frac{2\phi}{\overline{q}}X + b(P_I + P_C)}{b + \frac{2\phi}{\overline{q}}} = \overline{P})P(\frac{\frac{2\phi}{\overline{q}}X + b(P_I + P_C)}{b + \frac{2\phi}{\overline{q}}} = \overline{P})$$

The first expression

$$P((X - P_C \leq b\overline{q} + 2\phi + P_I) | \frac{\frac{2\phi}{\overline{q}}X + b(P_I + P_C)}{b + \frac{2\phi}{\overline{q}}} = \overline{P})$$

can be calculated as follows:

$$\frac{\frac{2\phi}{\overline{q}}X + b(P_I + P_C)}{b + \frac{2\phi}{\overline{q}}} = \overline{P} \Rightarrow X = \frac{(b + \frac{2\phi}{\overline{q}})\overline{P} - bP_I - bP_C}{\frac{2\phi\overline{P}}{\overline{q}}}$$

Substituting the result back one gets:

$$P((X - P_C \leq b\overline{q} + 2\phi + P_I) | \frac{\frac{2\phi}{\overline{q}}X + b(P_I + P_C)}{b + \frac{2\phi}{\overline{q}}} = \overline{P}) = P(P_C \geq \frac{(b + \frac{2\phi}{\overline{q}})(\overline{P} - P_I) - 2b\phi}{b + \frac{2\phi}{\overline{q}}})$$

$$= 1 - \mathrm{CDF}_{P_C}(\frac{(b + \frac{2\phi}{\overline{q}})(\overline{P} - P_I) - 2b\phi}{b + \frac{2\phi}{\overline{q}}})$$

Furthermore,

$$P(\frac{\frac{2\phi}{\overline{q}}X + b(P_I + P_C)}{b + \frac{2\phi}{\overline{q}}} = \overline{P}) = P(\frac{2\phi}{\overline{q}}X + bP_C = \overline{P}(\frac{2\phi}{\overline{q}} + b) - bP_I)$$

The above probability is the PDF of the sum of two lognormal variables which does not have an analytical solution. There are two strategies to approximate the pdf of this sum. First way is to use some approximations such as Fenton-Wilkinson (FW). These methods are proposed in the literature to give the distribution of the sum as a new lognormal variable after matching the first and second moments. According to the FW method (μ_Z, σ_Z) the mean and variance of new variable Z (the sum of two) are given by:

$$u_1 = \sum_{i=1}^{2} e^{m_i + \frac{\sigma_i^2}{2}}$$

$$u_2 = \sum_{i=1}^{2} e^{2m_i + 2\sigma_i^2}$$

$$m_Z = 2ln(u_1) - \frac{1}{2}ln(u_2)$$

$$\sigma_Z^2 = ln(u_2) - 2ln(u_1)$$

The second method is proposed by Gao, Xu, and Ye (2009). Their method gives the exact joint PDF of (e^{X_1}, e^{X_2}) as:

$$f(u,v) = \frac{h(u,v)}{2\phi\sigma_1\sigma_2\sqrt{(1-\rho^2)}uv}$$
$$h(u,v) = exp\left\{-\frac{1}{2(1-\rho^2)}[(\frac{ln(u)-\mu_1}{\sigma_1})^2 - 2\rho\frac{(ln(u)-\mu_1)(ln(v)-\mu_2)}{\sigma_1\sigma_2} + (\frac{ln(v)-\mu_2}{\sigma_2})^2]\right\}$$

Although this method provides the exact PDF of $(X_1, X_2$, we need to have the sum of probabilities over an interval since we want to have the probability of $X_1 + X_2 = Z$ which is equivalent to calculating the probability of $g(z) = \int_0^Z f(X_1, Z - X_1)dX_1$. Unfortunately, Gao, Xu, and Ye (2009) only provide such probabilities for tail distribution of the sum. One therefore needs to estimate the probability of $g(z)$ using some numerical integration methods.

Similarly, $X - P_C > b\overline{q} + 2\phi + P_I$ implies production in the binding region and the related event defined as

$$E_2 = \left\{(X, P_C) : (X - P_C > b\overline{q} + 2\phi + P_I) \wedge (X - b\overline{q} = \overline{P})\right\} \Rightarrow$$
$$P(E2) = P((X - P_C > b\overline{q} + 2\phi + P_I | X = \overline{q}) + \overline{P})P(X = \overline{q} + \overline{P})$$

The second probability is easy to calculate using the formula for conditional probability of bi-variate normal distribution.

VII. Applications

A. Optimal Hedging

Our model characterizes the dynamics of input/output price in an industry like refinery or airline. The interesting feature is the regime-switching behavior of the wedge between input and output prices because of capacity constraints. The model together with the calibrated demand process can be used to determine the optimal hedging policy of firms. Take the example of a

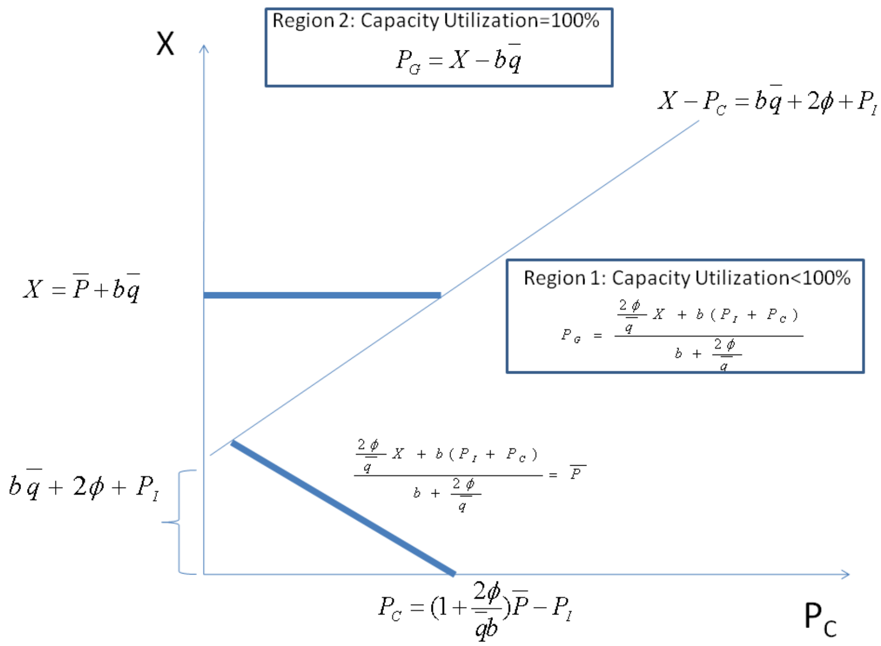

Figure 15. Event Space for Correlated Variables and Capacity Frictions

refinery. If demand is weak, input and output prices follow almost the same path with a fixed crack spread. The producer does not need any hedge since movements of input/output offset each other and leave a fixed margin for the producer. On the other hand, if demand is strong and industry is going to producer at boundary, output and input prices become independent and therefore crack spreads move stochastically which implies randomness in the cash flow of refinery.

Using the dynamics of demand process, a producer can decide whether a hedge of input or output is necessary and then calculate the optimal hedge policy.

B. Contingent Claim Pricing

C. Valuation

The dynamics of demand under risk-neutral measure can be used to estimate the value of a refinery or financial contracts such as futures or options on gasoline. To calculate the expected cash flow of refinery under risk-neutral measure, one needs to know the distribution of profit margin as well the optimal production rate. Forward contracts of gasoline and crude provide expectations regarding future values of these variables. Dynamics of X can be used to calculate optimal production rate and finally the expected cash flow.

References

Blair, B, and J 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, 2000, Sticky Prices, Inventories, and Market Power in Wholesale Gasoline Markets, Working Paper.

Buguk, C, D Hudson, and T Hanson, 2003, Price Volatility Spillover in Agricultural Markets: An Examination of U.S. Catfish Markets, *Journal of Agricultural and Resource Economics* 28, 86–89.

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

Considine, T. J., 1997, Inventories under Joint Production: An Empirical Analysis of Petroleum Refining, *The Review of Economics and Statistics* 79, 493–502.

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

Dempster, M.A.H, Eleva Medova, and Ke Tang, 2008, Long term spread option valuation and hedging, *Journal of Banking and Finance* 32, 2530–2540.

Gao, Xin, Hong Xu, and Dong Ye, 2009, Asymptotic Behavior of Tail Density for Sum of Correlated Lognormal Variables, *Economic Letters* 98, 229–234.

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.

Kuper, H G, and Tigran Poghosyan, 2008, Non-linear price transmission between gasoline prices and crude oil prices, Working Paper.

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