A Stochastic Supply/Demand Model for Storable Commodity Prices

Ali Bashiri and Yuri Lawryshyn

June 27, 2017

Abstract

We develop a two-factor mean reverting stochastic model for forecasting storable commodity prices and valuing commodity derivatives. We define a variable called "normalized excess supply" based on the observable production rate, consumption rate, and inventory levels of the commodity. Our analysis indicates a strong inverse correlation between normalized excess supply and crude oil spot and futures prices. Our first stochastic factor is normalized excess supply, following a mean reverting process. The second stochastic factor is the deviation of prices from a mean level determined by normalized excess supply. Moreover, we develop valuation models for futures and options on futures contracts based on the two factor model. We apply this model to crude oil prices from 1995 to 2017 via a Kalman filter. We perform out-of-sample tests for forecasting spot prices. Additionally, we develop a scenario analysis framework for incorporating variant views of future supply, demand, and inventories in risk management and valuation of commodity related financial products. We show the utility of the model for investment management professionals and risk managers aiming to incorporate macroeconomic conditions in valuation and risk management endeavors.

List of Figures

List of Tables

Contents

1 Introduction

Commodities are an integral part of the global economy. Produced and consumed globally, commodities have far-reaching impacts on economies and policies of importing and exporting nations. For instance, Hamilton (2008) and Morana (2013) demonstrate that oil price increase have preceded nine out of the ten most recent US recessions. Hamilton (2011) and Cuñado and de Gracia (2003) show the detrimental effects of oil price shocks on US and European Economies. For the case of copper, Ebert and La Menza (2015) demonstrate the importance of copper to the Chilean economy and its sustainable development as the largest producer of copper. The importance of commodities to national security has prompted governments and financial markets to pay close attention to commodities' price levels. Furthermore, commodity pricing plays an important role in real options valuation. Kobari et al. (2014) demonstrate the importance of oil prices in the real options valuation, operational decisions, and the expansion rate of Alberta oil sands projects. Bastian-Pinto et al. (2009) demonstrate the impact of sugar and ethanol price processes on the real option valuations of ethanol production from sugarcane. Brandão et al. (2013) show the significance of the soybean and castor bean price processes for real options valuation of managerial flexibility embedded in a biodiesel plant.

Moreover, in recent decades, commodities have been widely used as investment vehicles for financial institutions. Irwin and Sanders (2011) used data from Barclays Capital to demonstrate a surge of investment in commodity index products within the US and non-US markets from \$50 billion in 2004 to \$300 billion in 2010. In another study, Juvenal and Petrella (2015) estimate that assets allocated to commodity index trading had increased from $13 billion in 2004 to $260 billion in 2008. The importance and widespread use of commodities has led to growth of both the variety, and trading volume, of financial products related to commodities. According to statistics published by the Bank for International Settlements (BIS, 2015), the notional value of over-the-counter commodities' derivatives increased from $415 billion in 1998 to over $2 trillion in 2015. The increase in investment in commodity index products is attributed to the global liquidity increase. Beckmann et al. (2014) find a strong and time varying effect of global liquidity on commodity prices. Ratti and Vespignani (2015) find co-integration between M2 liquidity in BRIC and G3 economies and commodity prices with causality flowing from liquidity to commodity prices. The increase in appeal for commodity related investment stems from commodities' equity-like returns, protection against inflation, diversification, and impact on portfolios' risk and rewards (Geman, 2005; Gorton and Rouwenhorst, 2006; Belousova and Dorfleitner, 2012). The increase in investment in commodities, creates unprecedented scenarios with implications for investment management and project valuations, requiring a better understanding of price dynamics to manage these scenarios.

Many factors contribute to the pricing of commodities and related financial products. Commodity prices are a result of interactions between supply, demand, and other factors such as macroeconomic factors. Randomness and unpredictability of these factors result in prices exhibiting a stochastic behavior. The majority of previous literature has either focused on modeling the stochasticity of the prices or finding causal links between commodities' prices and a range of commodity specific and economic factors.

Stochasticity of commodity prices is addressed in various frameworks. A classical approach to modeling price dynamics is modeling the futures prices, unobservable convenience yield, and in some cases, interest rates, as stochastic factors. Gibson and Schwartz (1990), Schwartz (1997), Casassus and Collin-Dufresne (2005), Paschke and Prokopczuk (2010), Liu and Tang (2011), Chen and Insley (2012), Mirantes et al. (2013), and Lai and Mellios (2016) are some recent examples of literature focused on modeling stochasticity of commodity prices. Schwartz (1997) is a seminal work on modeling mean reversion and stochasticity of commodity prices, convenience yield, and interest rates.

Another approach to modeling commodity prices is to assume stochasticity in volatility of the prices as well as commodity spot prices. This approach stems from the view that volatility of commodity prices is stochastic. Specifically, these models try to match the volatility smile observed in the commodity options market. Hikspoors and Jaimungal (2008), Trolle and Schwartz (2009), Arismendi et al. (2016) are some of the recent works focused on modeling commodity prices as well as the volatility of the prices in a stochastic framework.

Another sector of related literature is focused on analyzing the impacts of commodity specific and economic factors on commodity prices and vice versa. Commodity specific factors refer to factors directly impacting the production and consumption of commodities ranging from supply/demand and inventory data to geopolitical and socioeconomic factors. Güntner (2014) analyzes the impact of oil prices and demand on oil production within OPEC and non-OPEC producers. Bu (2014) shows that weekly inventory data published by the U.S. Energy Information Agency (EIA) has a significant impact on oil prices. Tilton et al. (2012) classify commodity market participants as investors and consumers and conclude that investor demand, consumer demand, and producer supply play an important role in price discovery in copper markets. Elder et al. (2012) analyze the impact of macroeconomic news on metal futures markets and conclude that the response of markets to economic news is both swift and significant with unexpected improvements in the economy to have a negative impact on gold and silver prices, and a positive impact on copper prices.

Economic factors impact commodity prices by changing the expected supply and demand for commodities. Hong and Yogo (2012) find a strong correlation between open interest in commodity futures markets and macroeconomic activities. Casassus et al. (2013) demonstrate an economic link between related commodities' prices based on production, complimentary, and substitution relationships between said commodities. Examples of such commodities are lead, tin, zinc and copper, which are commonly used in alloy production.

Moreover, global economic activity is expected to have an impact on global energy demand. Kaminski (2014) classifies the North American energy market's important determinants as physical, financial, and socioeconomic layers, and measures each determinant's impact on the market. Stefanski (2014) concludes that the structural changes in the economies of developing countries account for the majority of price appreciations in energy markets.

Routledge et al.'s (2000) study is an influential theoretical work on modeling term structure of forward prices in equilibrium. In this work, the impacts of changes in inventory, net demand shocks, and convenience yield on forward prices are discussed. They show that the value of the convenience yield is impacted by interactions among supply, demand, and inventory.

Employing an econometric approach, Kilian (2014) demonstrates the importance of supply, demand, and inventory levels on commodity prices utilizing a vector autoregressive model. Kilian and Lee (2014) analyze the impact of speculative demand shocks on oil prices. Kilian (2014) demonstrates that supply and demand levels are integral parts of oil price dynamics. Baumeister and Kilian (2016) create a vintage dataset consisting of global oil production, US and OECD oil inventory levels and other factors, and demonstrate improved forecasting capabilities compared to similar vector autoregressive models.

Sidebottom et al. (2011) attribute increases in copper theft incidents to price increases in response to an excess global demand for the metal. Geman and Smith (2013) analyze and model the role of inventory and relative scarcity in the LME metals market based on the theory of storage. They find prices and price volatility to be higher in periods with lower inventory.

In summary, the literature demonstrates strong evidence of stochasticity of commodities' prices, convenience yield, and volatility. Moreover, the literature also suggests strong explanatory power with respect to commodity prices of production, consumption, and inventory levels.

In theory, commodity prices' dynamics should resemble mean reversion in equilibrium. In equilibrium, forces of supply and demand are in balance and, in absence of external influences, the prices should not change. An increase in demand would push the equilibrium price higher, resulting in an increase in supply as the market becomes attractive to new suppliers. The excess supply will in turn balance the initial increase in demand and reduce or reverse the upward pressure on prices. In the opposite scenario, a decrease in demand would create a glut in the market and push prices downward. A decrease in prices would pressure producers with higher costs to halt production, resulting in a reduction in available supply in the market. Consequently, the reduction in supply would decrease, or reverse, the downward pressure on prices. This balancing act between supply and demand is the fundamental force resulting in the mean reversion of prices.

In this work, we propose a two factor stochastic model. The first factor is "normalized excess supply", a factor incorporating the supply and demand for the commodity. And the second factor is the deviations of commodity's price from the mean level set by supply and demand conditions. This model is unique in that it proposes a mean reverting factor related to supply, demand, and inventory levels and measures the impact of changes in these variables on the commodity prices. Moreover, the model is calibrated not only on the commodity prices, but also on the observed production, consumption, and inventory levels. This model is then applied to various commodities ranging from oil to base metals. Following the Schwartz (1997) framework, since the true spot price, supply/demand, and inventory data of commodities are uncertain and unobservable, we use oil futures prices, observed production rate, consumption rate, and inventory data for model calibration. Specifically, the model is put in a state-space form and a Kalman filter is applied to estimate the model parameters and values of the state variables. As far as we are aware, no other stochastic models that include macroeconomic factors have yet been proposed. Our work is a pioneering effort within the newly established realm of "Quantamental Models", combining quantitative and fundamental analysis.

We apply the model to commodity market data from 1995 to present time. For crude oil, we use the West Texas Intermediate (WTI) futures data on a weekly basis as well as data on production rate, consumption rate, and inventory published by the Energy Intelligence Group (EIG). The base metals included in our analysis are copper, aluminum, nickel, zinc, lead, and tin. We use weekly futures data, consumption and production rates published by the World Bureau of Metals Statistics (WBMS), and data of available inventory at the Commodity Exchange (COMEX), the London Metal Exchange (LME), and the Shanghai Futures Exchange (SHFE).

The remainder of this work is organized as follows. Section 2 explains the economic rationale for the model. In Section 3, the forecasting model is explained and reviewed. In Section 4, the data is presented and described. Section 5 presents results of the model calibration and forecasting. Finally, Section 6 presents concluding remarks.

2 Economic Rationale

This section is dedicated to presenting the economic rationale for the proposed model. As previously stated, in equilibrium, commodity price dynamics are assumed to be mean reverting. An increase in demand leads to price appreciation. Price appreciation leads to a supply increase by allowing higher cost producers to enter the market. Increase in supply balances the excess demand and reduces or reverses the price appreciation. Conversely, a decrease in demand results in price depreciation. Price decrease results in losses for producers with higher costs and forces them to halt production, leading to reduced supply in the market. Reduction in supply balances the reduction in demand and creates an upward pressure on prices. In other words, price falls are generally followed by price increases and price increases are generally followed by price decreases due to the balancing act between supply and demand.

In this work, we propose a different mean reverting model for commodities. We postulate that commodity spot prices follow a mean reverting process with the difference that the mean reversion level is non-constant and dependent on market conditions regarding supply, demand, and inventory levels as well as other factors such as GDP and inflation. Chiang et al. (2015) conclude that oil prices have a statistically significant and economical relationship with real GDP. Furthermore, inflation leads to devaluation of currencies. The real value of a given commodity and costs associated with its production do not change as a unit of currency decreases or increases. This leads to an increase in the nominal value of a unit of commodity as inflation devalues the currency. Szymanowska et al. (2014) find evidence of the sizeable impact of inflation, as well other factors, on spot and term risk premia for a portfolio of seven commonly used commodities ranging from energy to agricultural products and industrial metals. Moreover, commodity returns are shown to be positively correlated with inflation (Greer, 2000; Erb and Harvey, 2006; Gorton and Rouwenhorst, 2006).

Finally, we believe commodities are impacted by observable supply, demand, and inventory levels. A positive demand shock (negative supply shock) would increase the price of the commodity while a negative demand shock (positive supply shock) would decrease the price. However, it is not the absolute demand or supply shock that would impact the price but the relative value of the shock. For instance, if an increase in demand is immediately met with an increase in supply, this shift would not contribute to scarcity of the commodity and price increases. Moreover, inventories act as a buffer for the markets by absorbing excess supply and offsetting excess demand. However, it should be noted that inventories only provide buffering capacity for storable commodities (see Routledge et al. (2000); Carlson et al. (2007); Sockin and Xiong (2015) for discussion on impacts of inventory on storable commodities). Moreover, inventories cease to provide buffering protection for positive supply shocks at near maximum capacity. This is due to the limited capacity for storage and not being able to mix different qualities of a commodity such as crude oil in the same container (Kaminski, 2014). We propose a ratio called "normalized excess supply", defined as

$$q = \frac{Supply - Demand}{Demand} \tag{2.1}$$

to capture the impact of supply, demand, and inventory on commodity spot prices. We believe that normalized excess supply plays an important role in price setting and understanding its dynamics is essential. We put forward that normalized excess supply should exhibit mean reverting properties. In the long term, there should be a balance between supply, demand, and inventory levels. Supply and inventory levels should always be in line with demand. Therefore, the ratio of excess total supply to demand should be mean reverting in the long term. Deviation from this long term mean level would have an impact on prices. With increasing total supply, this ratio would increase and in turn, prices would decrease. We formulate the model such that this ratio impacts the mean reversion level of spot prices through a linear function

$$\mu = aq + b \tag{2.2}$$

where "a" and "b" are constants that can be fitted to data. For instance, with increasing supply, q would increase. Producers and market participants are expected to store the commodity at this point. It is then expected that the total return on prices would decrease providing incentive for participants to hold rather than sell the commodity. The increasing abundance of the commodity could even reduce the prices if the magnitude of the supply increase is significant. In the next section, details of the proposed model are presented.

3 The Model

This section outlines the proposed model and derivatives' valuations. The logarithm of spot prices, X, is modeled as combination of two mean reverting processes. The two factors of this model are the normalized excess supply of the commodity, q, and the deviations of X from the mean level of log-prices, defined as ε. The two factors are modeled as the following joint stochastic processes

$$X(t) = aq(t) + \varepsilon(t) + b, \tag{3.1}$$

$$dq(t) = \kappa_1(\theta_1 - q(t))dt + \sigma_1 dz_1, \tag{3.2}$$

$$d\varepsilon(t) = -\kappa_2 \varepsilon(t)dt + \sigma_2 dz_2, \tag{3.3}$$

where,

- X is the log of spot prices, S,
- κ_1 and κ_2 are the rates of mean reversion,
- θ_1 is the level of mean reversion,
- a and b are constants determining the relationship between level of prices and "q",
- σ_1 and σ_2 are the volatilities,

– z_1 and z_2 are standard Brownian motions,

and the two standard Brownian motions are correlated as

$$dz_1 dz_2 = \rho dt, \tag{3.4}$$

where ρ is the correlation between the two motions.

Equation 3.1 demonstrates the log-price process. Equation 3.2 describes the normalized excess supply (q) as an OU mean reverting stochastic process. For the purpose of this analysis, q is calculated as

$$q(t) = \frac{I(t-1) + P(t) - C(t)}{C(t)}, \tag{3.5}$$

where $I(t-1)$ represents inventory level at the beginning of the period, $P(t)$ is the total production during the period, and $C(t)$ is the total consumption throughout the period. The combination "$I(t-1)+P(t)$" is seen as the total supply available while "$C(t)$" is seen as the total demand. The combination "$aq(t)+b$" represents the mean level of log-prices. Deviations of log-prices from this mean level are defined as $\varepsilon(t)$ represented in equation 3.3. $\varepsilon(t)$ is defined as a zero-mean level mean reverting process. Essentially, log-prices are defined as a mean reverting process with the mean level defined in equation 2.2.

In the proposed model, X is modeled as a 2-dimensional mean reverting system, where the value of q determines the mean reversion level for X. A 2-dimensional spring system could further help in understanding the dynamics of this system as demonstrated in figure 3.1. Normalized excess supply is modeled as a mean reverting process, similar to a spring as shown in figure 3.1, the gray spring. Furthermore, the ε is also modeled as a mean reverting process, the blue spring, such that it's mean reversion (equilibrium) level is zero. In a 2-D spring system, the q spring is transformed by $aq+b$ such that it lies on the mean reversion level of log-prices.

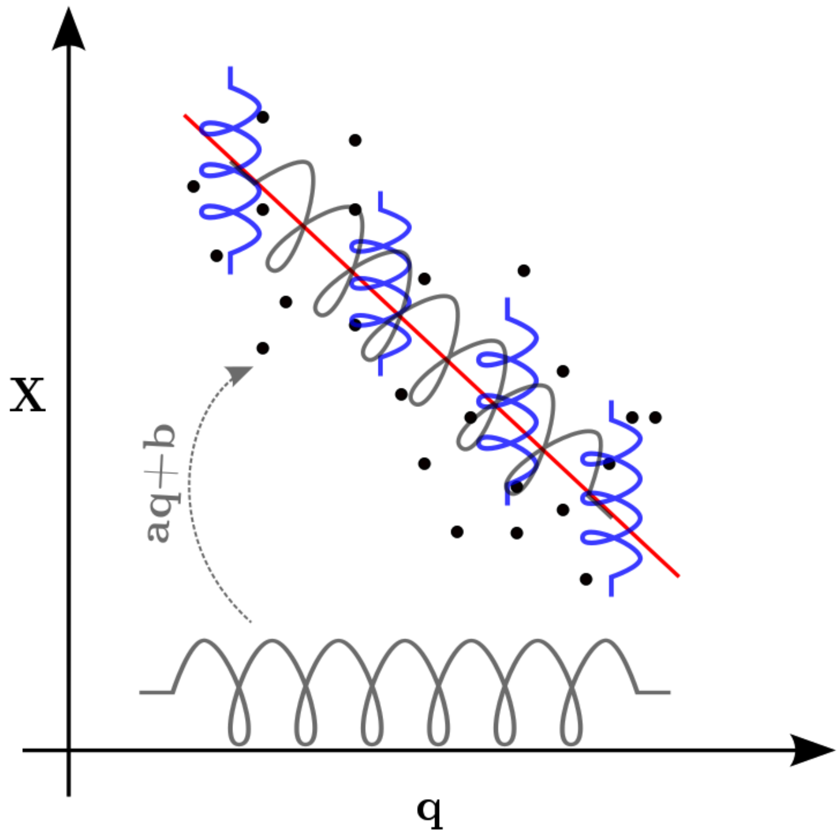


Figure 3.1: Two Dimensional Spring Mean Reverting System. The q spring (gray) is transformed along $aq+b$ and is combined with the ε spring (blue) to model the spot prices, X.

Under the risk-neutral framework, applying the Girsanov's theorem, the model can be presented as

$$X(t) = aq(t) + \varepsilon(t) + b, \tag{3.6}$$

$$dq(t) = [\kappa_1 (\theta_1 - q(t)) - \lambda_1] dt + \sigma_1 d\hat{z}_1, \tag{3.7}$$

$$d\varepsilon(t) = (-\kappa_2 \varepsilon(t) - \lambda_2) dt + \sigma_2 d\hat{z}_2, \tag{3.8}$$

$$d\hat{z}_1 d\hat{z}_2 = \rho dt, \tag{3.9}$$

where,

– $\hat{z}_1$ and $\hat{z}_2$ are standard Brownian motions,

– λ_1 and λ_2 are the market prices of risks.

This joint distribution for q and ε can be then presented as

$$\begin{pmatrix} q(t) \\ \varepsilon(t) \end{pmatrix} \sim \mathscr{N}\left(\begin{bmatrix} \mu_q(t) \\ \mu_\varepsilon(t) \end{bmatrix}, \begin{bmatrix} \sigma^2_{q(t)} & \sigma_{q(t)\varepsilon(t)} \\ \sigma_{q(t)\varepsilon(t)} & \sigma^2_{\varepsilon(t)} \end{bmatrix} \right), \tag{3.10}$$

where $\mathscr{N}(.)$ represents a bivariate normal distribution. The terms $\mu_{q(t)}$, $\mu_{\varepsilon(t)}$, $\sigma^2_{q(t)}$, $\sigma^2_{\varepsilon(t)}$, and $\sigma_{q(t)\varepsilon(t)}$ are derived in Appendix 7.1 and presented below

$$\mu_{q(t)} = \mathrm{e}^{-\kappa_1 t} q(0) + \theta_1 \left(1 - \mathrm{e}^{-\kappa_1 t}\right), \tag{3.11}$$

$$\mu_{\varepsilon(t)} = \mathrm{e}^{-\kappa_2 t} \varepsilon(0), \tag{3.12}$$

$$\sigma^2_{q(t)} = \frac{1}{2} \frac{\sigma_1^2 \left(1 - \mathrm{e}^{-2\kappa_1 t}\right)}{\kappa_1}, \tag{3.13}$$

$$\sigma^2_{\varepsilon(t)} = \frac{1}{2} \frac{\sigma_2^2 \left(1 - \mathrm{e}^{-2\kappa_2 t}\right)}{\kappa_2}, \tag{3.14}$$

$$\sigma_{q(t)\varepsilon(t)} = -\frac{\sigma_1 \sigma_2 \rho \left(\mathrm{e}^{-t(\kappa_1+\kappa_2)} - 1\right)}{\kappa_1 + \kappa_2}. \tag{3.15}$$

The log-spot prices, X, are then normally distributed, $\mathscr{N}(\mu_{X(t)}, \sigma^2_{X(t)})$, with

$$\mu_{X(t)} = a\left(\left(1 - \mathrm{e}^{-t\kappa_1}\right)\theta_1 + \mathrm{e}^{-t\kappa_1} q(0)\right) + \mathrm{e}^{-t\kappa_2}\varepsilon(0) + b, \tag{3.16}$$

and

$$\sigma^2_{X(t)} = \frac{a^2 {\sigma_1}^2 \left(1 - \mathrm{e}^{-2t\kappa_1}\right)}{2\kappa_1} + \frac{{\sigma_2}^2 \left(1 - \mathrm{e}^{-2t\kappa_2}\right)}{2\kappa_2} + 2\frac{a\sigma_1 \sigma_2 \rho \left(1 - \mathrm{e}^{-t(\kappa_1+\kappa_2)}\right)}{\kappa_1 + \kappa_2}. \tag{3.17}$$

Applying the Feynman-Kac theorem, any derivative under this model should satisfy the following partial differential equation for undiscounted derivatives G

$$0 = G_t + [\kappa_1(\theta_1 - q) - \lambda_1]G_q + (-\kappa_2\varepsilon - \lambda_2)G_\varepsilon + \frac{1}{2}\sigma_1^2 G_{qq} + \frac{1}{2}\sigma_2^2 G_{\varepsilon\varepsilon} + \sigma_1\sigma_2\rho G_{q\varepsilon}. \tag{3.18}$$

As previously mentioned, we use futures data for calibration of this model. There are two popualar methods for pricing futures contracts, the expectation method and the no-arbitrage method. Under the expectation method, for a futures contract, the PDE in equation 3.18 is subject to terminal condition $F(T,0) = S_T$. A futures contract's value at time t with maturity $\tau = T - t$, denoted as $F(t,\tau)$, is obtained as the expectation of the spot price at maturity. Given the log-normal distribution of spot price, we have

$$\begin{aligned} F(t,\tau) &= E^Q[S(T)] \\ &= \mathrm{e}^{\mu^Q_{X(T)} + \frac{1}{2}\sigma^2_{X(T)}}. \end{aligned} \tag{3.19}$$

where $\mu^Q_{X(T)}$ is the mean of log-spot price under the risk-neutral measure derived as

$$\mu^Q_{X(T)} = a\left(\left(1-\mathrm{e}^{-(T-t)\kappa_1}\right)\hat{\theta}_1 + \mathrm{e}^{-(T-t)\kappa_1}q(t)\right) - \frac{\lambda_2}{\kappa_2}\left(1-\mathrm{e}^{-(T-t)\kappa_2}\right) + \mathrm{e}^{-(T-t)\kappa_2}\varepsilon(t) + b, \tag{3.20}$$

with

$$\hat{\theta}_1 = \theta_1 - \frac{\lambda_1}{\kappa_1}. \tag{3.21}$$

Substituting for $\mu^Q_{X(T)}$ and $\sigma^2_{X(T)}$ in equation 3.19 and reorganizing the equation, we obtain

$$F(t,\tau) = \mathrm{e}^{\alpha(\tau)+\beta(\tau)q(t)+\gamma(\tau)\varepsilon(t)}, \tag{3.22}$$

where

$$\begin{aligned} \alpha(\tau) = {} & a\hat{\theta}_1\left(1-\mathrm{e}^{-\tau\kappa_1}\right) - \frac{\lambda_2}{\kappa_2}\left(1-\mathrm{e}^{-\tau\kappa_2}\right) + \frac{a^2{\sigma_1}^2\left(1-\mathrm{e}^{-2\,\tau\kappa_1}\right)}{4\kappa_1} + \frac{{\sigma_2}^2\left(1-\mathrm{e}^{-2\,\tau\kappa_2}\right)}{4\kappa_2} \\ & + \frac{a\sigma_1\,\sigma_2\,\rho\left(1-\mathrm{e}^{-\tau(\kappa_1+\kappa_2)}\right)}{\kappa_1+\kappa_2} + b, \end{aligned} \tag{3.23}$$

$$\beta(\tau) = \mathrm{e}^{-\kappa_1\tau}, \tag{3.24}$$

and

$$\gamma(\tau) = \mathrm{e}^{-\kappa_2\tau}. \tag{3.25}$$

It should be noted that under these assumptions, log-futures prices follow a normal distribution $\mathcal{N}(\mu_{F(t,\tau)}, \sigma^2_{F(t,\tau)})$ with

$$\mu_{F(t,\tau)} = \alpha(\tau) + \beta(\tau)\mu^Q_{q(T)} + \gamma(\tau)\mu^Q_{\varepsilon(T)}, \tag{3.26}$$

and

$$\sigma^2_{F(t,\tau)} = \beta(\tau)^2\sigma^2_{q(T)} + \gamma(\tau)^2\sigma^2_{\varepsilon(T)} + 2\beta(\tau)\gamma(\tau)\sigma_{q(T)\varepsilon(T)}. \tag{3.27}$$

Alternatively, following the no-arbitrage method, the futures price, $F(t,\tau)$, is derived as

$$F(t,\tau) = S_t\mathrm{e}^{(r-\delta)\tau}, \tag{3.28}$$

where r is the risk-free interest rate and δ is the net convenience yield for maturity τ. Under the no-arbitrage assumption, log-futures prices follow a similar distribution as the log-spot prices. Specifically, log-futures are normally distribution, $\mathscr{N}(\mu_{F(t,\tau)}, \sigma^2_{F(t,\tau)})$ with

$$\mu_{F(t,\tau)} = \mu_X + (r - \delta)\tau, \tag{3.29}$$

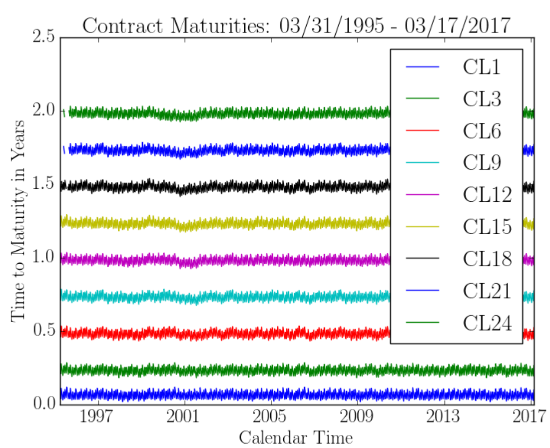
and

$$\sigma^2_{F(t,\tau)} = \sigma^2_X. \tag{3.30}$$

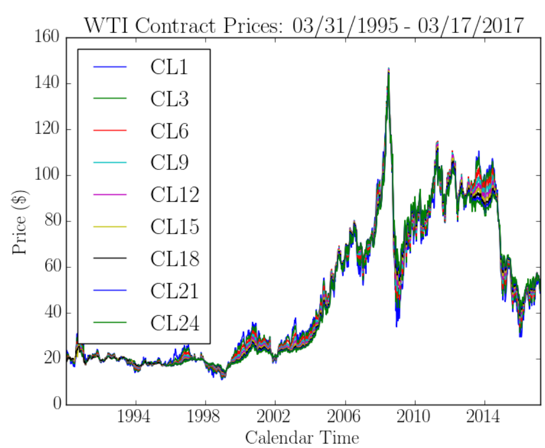
The no-arbitrage and expectation methods should yield consistent results. In theory, changes in the futures price $F(t,\tau)$ should be related to changes in the anticipated distribution of spot prices at maturity, not the current spot price. However, as suggested in Black (1976), in practice, the futures prices follow the current spot prices. Black (1976) suggests that changes in the cost of production, supply, and demand impact both futures and spot prices, resulting in high correlation between spot and future price changes. Furthermore, events such as the arrival of the production season only impacts spot prices. Brennan (1958) and Telser (1958) demonstrate that factors such as interest rates and cost of storage impact futures and spot prices. In light of our aim in this work to step out of the theoretical realm and model the observed spot-futures relationship, we favor the no-arbitrage method over the expectation method. In the next section, we present the data used for calibration.

4 Data

In this section, we apply the proposed model to weekly commodities data from March 1995 to March 2017. Specifically, crude oil, copper, aluminum , nickel, zinc, tin, and lead futures prices, production, consumption, and inventory data are used. Futures are available maturing every month of the year. We select eight futures contracts for calibration of the model, the 1st, 3rd, 6th, 9th, 12th, 15th, 18th, 21st, and 24th nearby contracts. These specific contracts are selected such that market views and expectations corresponding to the next 8 quarters. Our choice of time frame is consistent with economic forecasting standards. The selected contracts represent some of the most highly liquid available contracts. Figure 4.1a represents the maturity of the contracts used for calibration and WTI prices for sample.



(a) Futures' Maturity Data



(b) Futures' Price Data

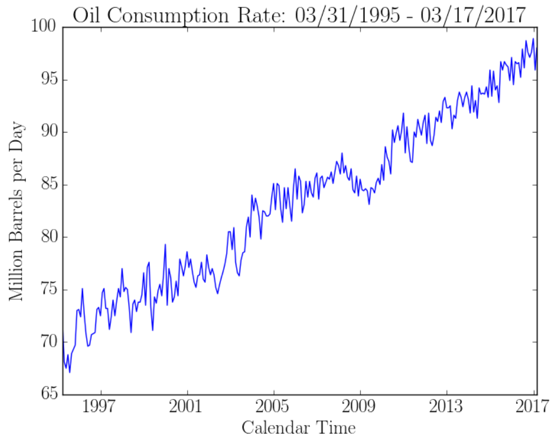
Figure 4.1: WTI Futures Contracts Weekly Information. Weekly West Texas Intermediate futures are used for calibration and estimations from March 1995 to March 2017.

Moreover, for supply, demand and inventory values, we use data published by international organizations dedicated to gathering and publishing of such data. Specifically, we use published reports by the Energy Intelligence Group (EIG) for crude oil and the World Bureau of Metals Statistics (WBMS) for base metals.

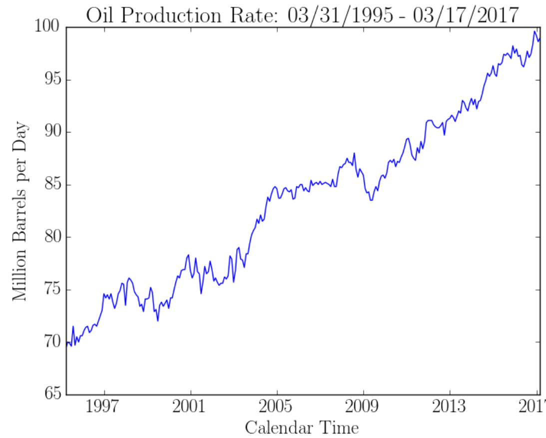
EIG publishes data on global supply and demand on a monthly basis. The inventory data is only available for the Organization for Economic Cooperation and Development (OECD) countries, which causes a discrepancy as there are other storage facilities outside of the domain of the OECD, mainly China, and non-OECD inventory data is not readily available. However, as demonstrated by Kilian and Murphy (2014), OECD inventory levels provide a good proxy for global oil inventory levels.

WBMS publishes global supply and demand data for base metals on a monthly basis. Furthermore, base metals' futures are mainly traded on three exchanges around the world, namely the Commodity Exchange (COMEX), the London Metal Exchange (LME), and the Shanghai Futures Exchange (SHFE). These exchanges are the main publicly accessible inventory depots for storage of base metals. The inventory statistics used in this paper are the cumulative available inventory at these exchanges.

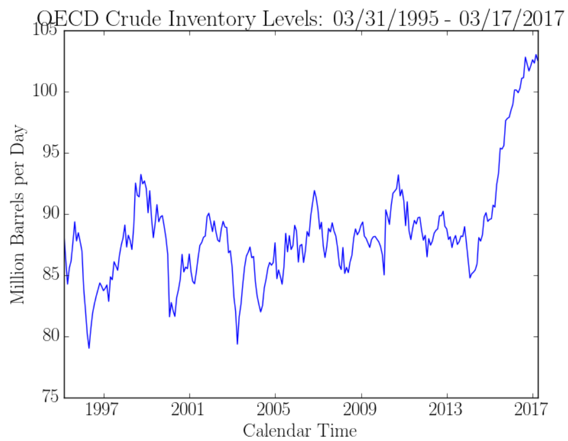
Figures 4.2a, 4.2b, and 4.2c demonstrate the monthly data on consumption, production, and inventory. Similar figures for copper, aluminum, nickel, zinc, lead, and tin are presented in the appendix 7.3.



(a) Global Crude Oil Consumption Rate



(b) Global Crude Oil Production Rate



(c) OECD Countries' Crude Oil Inventory

Figure 4.2: Crude Oil Consumption, Production, and Inventory Data. Monthly consumption, production, and inventory data are used for estimating normalized excess supply.

Moreover, we obtain U.S. Treasury Rates with maturities ranging from 1 month to 2 years published by the U.S. Federal Reserve. Then, appropriate interest rates for each specific derivative's maturity is calculated via cubic splines. The short term interest rate of less than 1 month is assumed constant at the 1 month maturity rate. Figure 4.3 demonstrates the yield rates for the period under study.

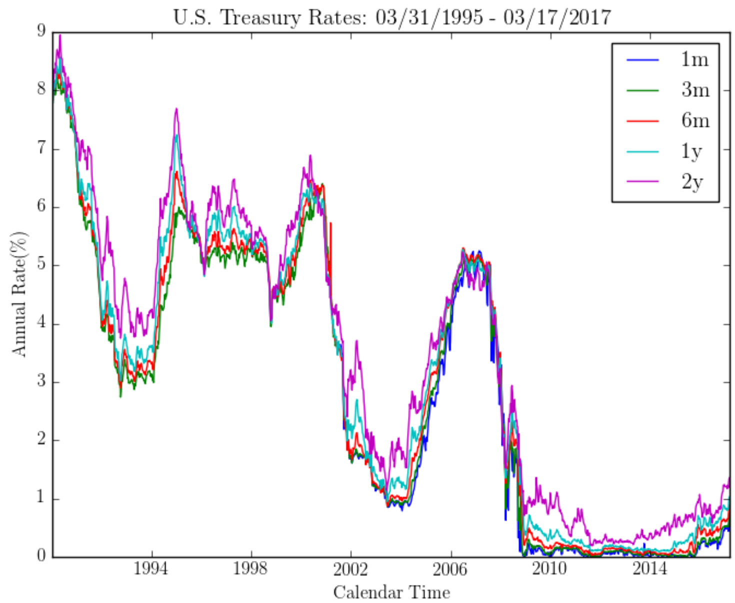


Figure 4.3: US Treasury Rates. T-Rates are used as a proxy for risk-free interest rate in calibration and analysis.

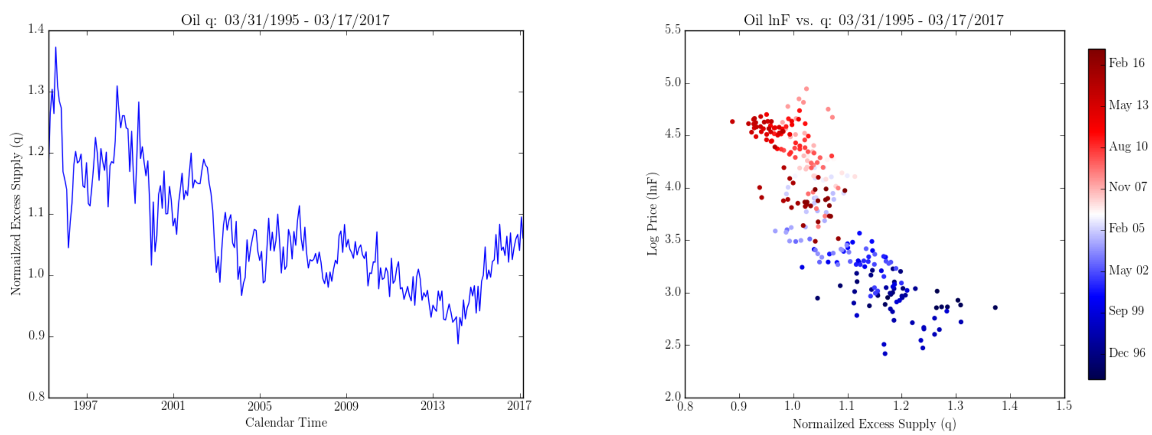
For the purpose of calibration, a Kalman filter process similar to that of Schwartz (1997) is utilized. The two underlying stochastic factors in the proposed model are normalized excess supply, q, and the deviation from mean level of log-prices, ε. Spot prices are unobservable for most commodities. Often, first nearby futures' prices are used as a proxy for the spot prices. Moreover, the published supply, demand, and inventory data are only estimates of the true unobservable variables. This makes straightforward calibration of the processes inaccurate. Hence, the model is put in a state-space form to account for unknown true values of the state variables. The two state variables are the q and ε. The Kalman filter is then applied to estimate the true value of the state variables' time series. Details of the Kalman filter are presented in Appendix 7.5.

5 Results

In this section, the results of data analysis, model calibration, and scenario analysis are presented. In section 5.1, we conduct statistical analysis on the constructed normalized excess supply (q) and deviations from mean (ε) for crude oil. In section 5.2, we presented a sliding window calibration analysis on crude oil futures. Calibration results are only presented for crude oil for the sake of brevity and preventing unnecessary repetition. We present results of calibration, in-sample estimation, and out-of-sample forecasting performance of the model. Additionally, we present out-of-sample performance of the Schwartz 2-factor model (Schwartz, 1997) for reference. Section 5.3 is dedicated to two case studies, analyzing the out-of-sample performance of the model during the 2008 and 2014 oil market crashes. In section 5.4 we present a scenario analysis framework based on our model. The framework is of value to risk managers, investment analysts, and traders.

5.1 Data Analysis

In this section, we analyze the statistical properties of the state variables' time series for crude oil. Figure 5.1 demonstrates the monthly normalized excess supply (q) and the relationship between log-prices and q for crude oil.



(a) Crude Oil Normalized Excess Supply "q" **(b)** Nearby Futures vs. Normalized Excess Supply

Figure 5.1: Crude Oil q and $\ln(F) vs. q$. This figure shows monthly normalized excess supply and first nearby future vs. "q" for crude oil. The relationship is significant with a negative correlation of 0.85.

For crude oil, as the data on supply and demand are rates of production and consumption per day, we convert the data to obtain total production and consumption for the month. As figure 5.1b demonstrates, there is a linear relationship between q and log-prices. This is consistent with the theoretical framework presented in section 3 and figure 3.1. Assuming the linear relationship $\mu = aq + b$ (eq. 2.2), we perform a linear regression between q and log-prices to obtain estimates of "a" and "b". We then subtract μ from the log-prices to obtain the deviations from the mean level. Figure 5.2 shows the resultant time series for the deviations, ε. As suggested in section 3, the deviations (ε) seem to follow a mean reverting process around a mean level of zero. Equivalent figures of the q and ε times series as well as $\ln(F) vs. q$ for copper, aluminum, nickel, zinc, lead, and tin are presented in appendix 7.4.

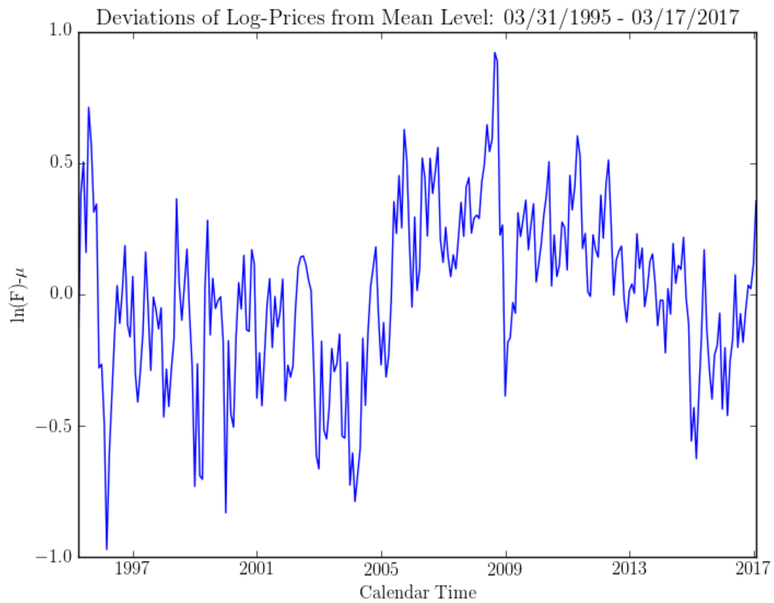


Figure 5.2: Crude Oil Price Deviations from Mean Level. ε seems to follow a mean reverting process around zero. The plot is generated for whole period of March 1995 to March 2017.

The next step is analyzing suitability of the data for the proposed model. We test the data for mean reversion assumptions of q and ε. Specifically, using Jarque-Bera (JB) test (Bera and Jarque, 1987), Augmented Dickey Fuller (ADF) test (Dickey and Fuller, 1979), variance ratio (VR) test, and the Generalized Hurst Exponent (GHE) test (Hurst, 1951; Mandelbrot, 2003), we examine the data for normality, stationarity, and mean reversion. The results are shown in table 5.1. The JB test suggests that the data are normally distributed. VR and ADF tests for ε suggest that the data series has some predictability and is stationary. For q time series, the VR test suggests predictability of the data, while the ADF test suggests that the time series is trend-stationary. Given that q should be stationary and mean reverting from an economics point of view, we deem the ADF test results not concerning. The GHE tests return values less than 0.5, suggesting that both time series are mean reverting. In summary, the theoretical model is generally appropriate for the data set.

Table 5.1: Statistical Tests on Data

	Jarque-Bera[1]	**Variance Ratio Test**[2]	**ADF Test**[3]	**Hurst Exponent Ratio**[4]
q	0	1	0, TS[5] = 1	0.33
ε	0	1	1	0.35

[1] JB test examines the null hypothesis that the data comes from a normal distribution with an unknown mean and variance. Value of "0" indicates failure to reject the null hypothesis.

[2] VR test assesses the null hypothesis of a random walk in a univariate time series. Value of 1 indicates rejection of the random-walk null in favor of the alternative.

[3] ADF test examines data for unit-root null hypothesis. Value of 0 indicates failure to reject the null hypothesis while a value of 1 indicates rejection of null hypothesis in favor of alternative.

[4] GHE test provides a ratio indicating whether a series is mean reverting, random walking or trending. Ratio values less than 0.5 suggest mean reversion, equal to 0.5 suggest random walk, while greater than 0.5 suggest trending time series.

[5] Trend-stationary test for q suggest that the data is trend stationary. That is, subtracting the mean value of the data from the data point transforms the time series into a stationary one.

5.2 Calibration

For the purpose of calibration, we follow the process set out in appendix 7.5. Specifically, we use two rolling windows and an expanding window frame works to calibrate and test the model and analyze the results. We implement three tests; One with a 3 year rolling window, one with a 5 year rolling window, and one with an expanding window. The windows move on monthly basis as shown in figure 5.3. These configurations resulted in respectively 250, 227, and 251 windows and instances of calibration for the 3 year, 5 year, and expanding windows alternatives.

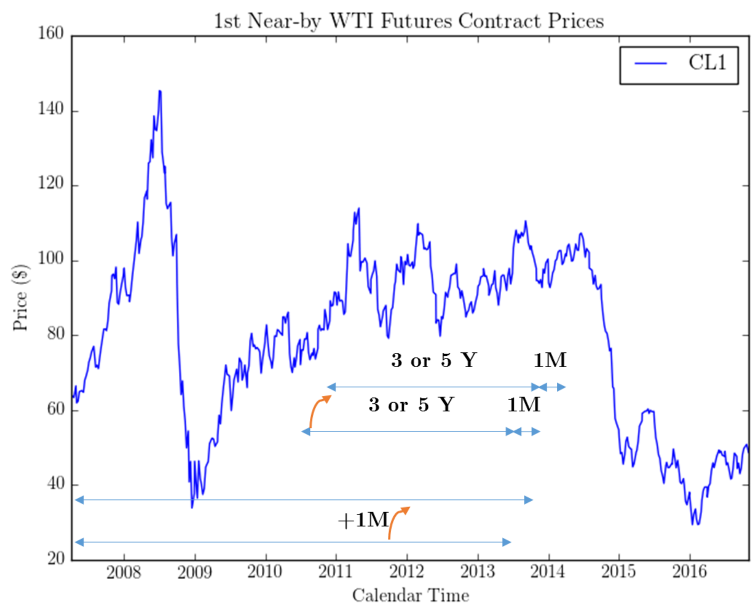


Figure 5.3: Rolling and Expanding Window Calibrations. Rolling windows of 3 and 5 years and an expanding window are used for calibration. The windows slide one month at a time. The frequency of data used for calibration is weekly.

One of the outputs of our calibration is the set of convenience yields of equation 3.28 associated with each contract, at each point in time. This convenience yield is the yield net of any costs associated with holding the commodity such as storage costs, insurance, and spoilage. Figure 5.4 demonstrates the annualized market implied net convenience yield. As expected, the convenience yield is more volatile in the short-term and plateaus as the term to maturity increases. This is due to market participants anticipating near-term supply/demand imbalances while the longer-term yields are dominated by costs of storing the commodity, resulting a smoothing effect at the longer maturity term end of the figure.

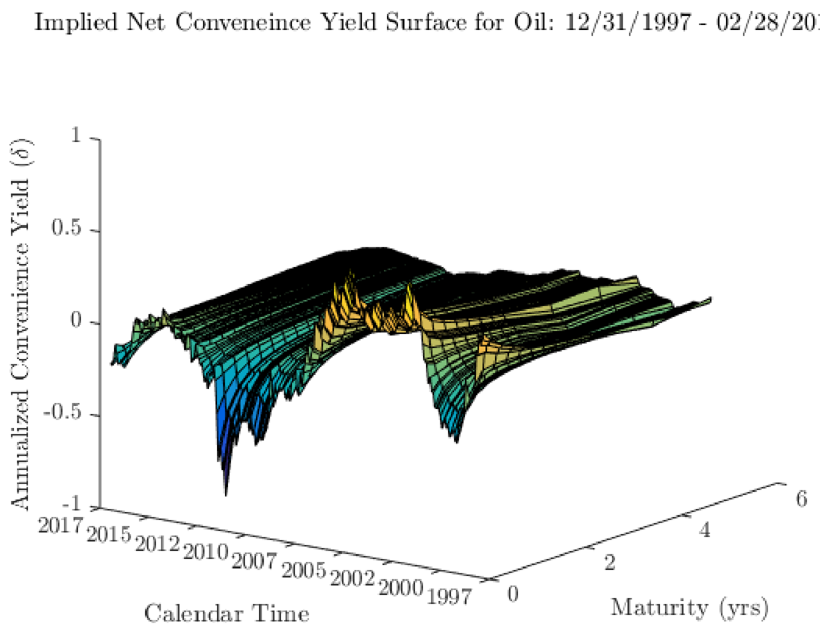
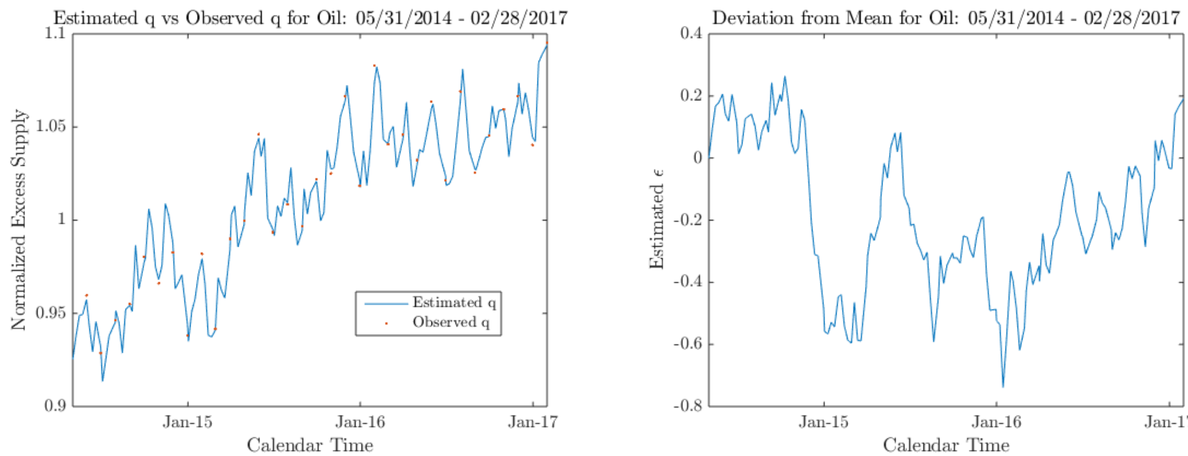


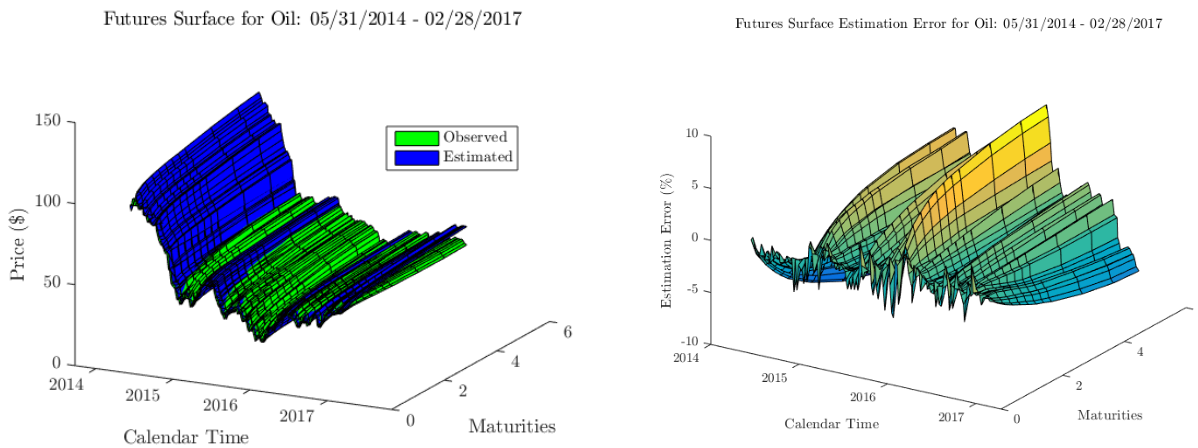
Figure 5.4: Estimate of Convenience Yields. Based on rolling window analysis of 3 years for 1995:2017

Some other outputs of the calibration process are the state variables, q and ε. Figures 5.5a and 5.5b demonstrate these estimations for the most recent of the calibration periods, May 2014 to February 2017. Figure 5.5a shows the closeness of the fit for normalized excess supply and that visually the estimated q matches the observed q very well. Figure 5.6 shows a sample estimation of the futures prices' surface. Specifically, figure 5.6a demonstrates the estimated and observed surface while figure 5.6b shows the percentage error in the estimation. The estimation error is less than 10%, indicating a close fit.



(a) Observed q vs. Estimated q **(b)** Estimated ε

Figure 5.5: Sample Estimation of State Variables q and ε. These estimations correspond to the latest 3 year window 2014:2017. The estimation of q produced low errors while the true ε is unobserved.

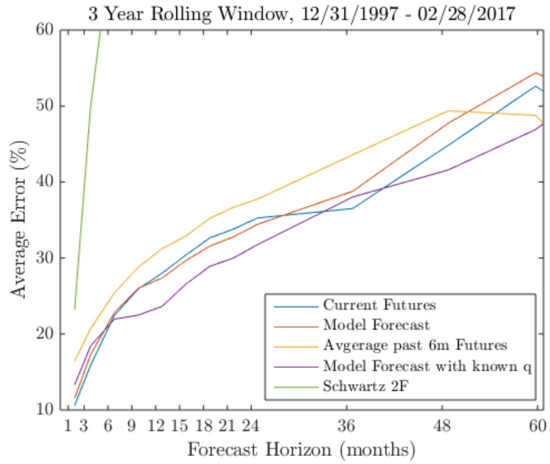


(a) Observed Futures vs Estimated Futures **(b)** Estimation Error

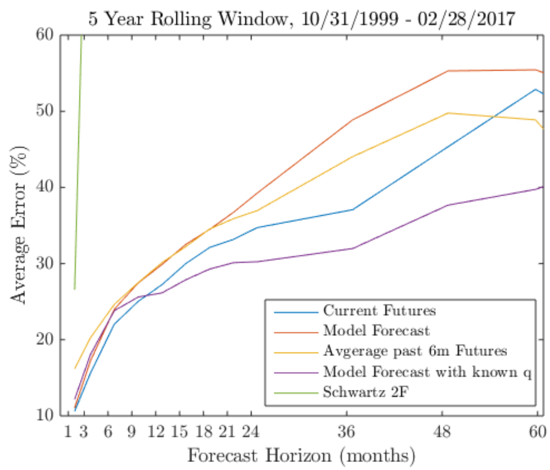
Figure 5.6: Sample Estimation Performance of Futures Surface. For the latest estimation period of 2014:2017. Errors are generally contained within 10% and increase with time-to-maturity.

Furthermore, we conduct out-of-sample tests of the model by forecasting the futures prices for maturities of up to 5 years. Forecasting the oil prices is a popular topic of research. As surveyed by Alquist et al. (2013), there are a fair number of papers in this area, however, it is hard to reconcile the conflicting evidence presented due to different estimation data, techniques, assumptions, and periods under study. However, the most common problem in the literature is the sensitivity of the results to the period under analysis and vanishing advantages when periods are expanded. In this regard, Alquist et al. (2013) reports the "no-change" forecast to outperform other models in both short and long term forecasting. Here, we conduct forecasting tests on various periods to present an overall measure of performance and present both the accuracy and precision of forecasts. Accuracy is defined as mean absolute percentage error and precision is defined as the standard deviation of the percentage errors. Figure 5.7 and table 5.2 show the performance of the out-of-sample forecast for the three calibration tests.

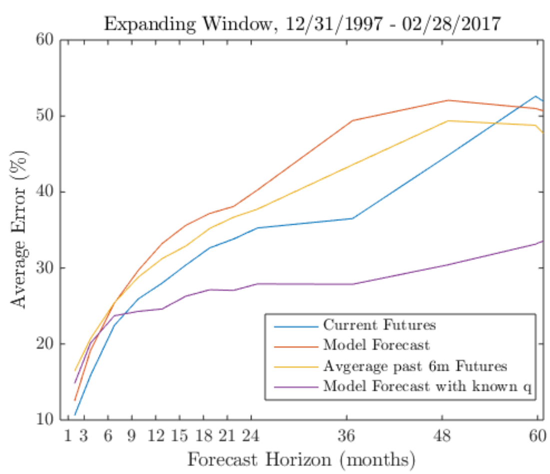
For this analysis, we forecast the prices as as an expectation of X in equation 3.1. Furthermore, the current futures prices are used as expectations of the market participants regarding the future spot prices. Given the high correlation between spot and futures prices, this benchmark is similar to the no-change benchmark. Also, as another benchmark, we have included results of the 2-factor Schwartz model (Schwartz, 1997). In order to measure the performance of the model over time, the forecasting test is conducted on all of the calibrated periods. Figure 5.7 demonstrates the mean absolute forecasting errors over the 1995-2017 period for various horizons. As demonstrated, the model's blind forecast performance is comparable on average to market participants' performance. However, it has lower standard error of forecast as presented in table 5.2. Additionally, the advantage of this model is allowing inclusion of variant views on supply, demand, and inventory levels. In order to evaluate the model's performance equipped with perfect estimates of supply, demand, and inventory, we allow for look-ahead to use actual future observed q values. Equipped with perfect estimates of q, the model outperforms the market. In summary, understanding the dynamics of q and being able to forecast values of q accurately, plays an important role in accurately forecasting the future prices. Comparison between the 3 year, 5 year, and expanding windows suggest that the blind model performs better in 3-year window test in the shorter term and performs better in the 60 months forecasting with expanding window calibration. Performance of the model equipped with q improves significantly with the expanding window calibration measured by both accuracy and precision. The Schwartz (1997) 2 factor model takes considerably longer to calibrate under equivalent conditions. Furthermore, the results are unsatisfactory due to low accuracy and precision. We have limited the Y-Axis in figures 5.7 to highlight the distinctions of different models and calibration setups.



(a) 3 year window calibration



(b) 5 year window calibration



(c) Expanding window calibration

Figure 5.7: Out-of-sample Forecasting Performance. Figures above demonstrate the mean absolute percentage error of forecasts under different calibration setups and time horizons.

Table 5.2: Out-of-sample Forecasting Performance

Maturity (months)	Current Futures	Average Futures	3 Year Windows: NOBS	3 Year Windows: Model	3 Year Windows: Model q	3 Year Windows: Schwartz 2F	5 Year Windows: NOBS	5 Year Windows: Model	5 Year Windows: Model q	Expanding Window: NOBS	Expanding Window: Model	Expanding Window: Model q
1	0.106 (0.094)	0.164 (0.204)	249	0.115 (0.098)	0.133 (0.117)	0.166 (0.162)	225	0.111 (0.100)	0.122 (0.116)	225	0.125 (0.099)	0.148 (0.127)
3	0.159 (0.186)	0.207 (0.249)	247	0.173 (0.171)	0.184 (0.175)	0.333 (0.370)	223	0.173 (0.197)	0.181 (0.196)	223	0.192 (0.166)	0.201 (0.190)
6	0.225 (0.287)	0.254 (0.276)	244	0.228 (0.252)	0.220 (0.237)	0.494 (0.558)	220	0.240 (0.293)	0.238 (0.285)	220	0.254 (0.231)	0.237 (0.229)
9	0.259 (0.265)	0.288 (0.270)	241	0.260 (0.253)	0.225 (0.235)	0.554 (0.559)	217	0.275 (0.270)	0.256 (0.256)	217	0.297 (0.242)	0.243 (0.206)
12	0.280 (0.250)	0.312 (0.293)	237	0.274 (0.247)	0.236 (0.239)	0.565 (0.584)	213	0.299 (0.248)	0.262 (0.232)	213	0.332 (0.25)	0.246 (0.200)
15	0.304 (0.268)	0.329 (0.317)	234	0.297 (0.289)	0.266 (0.286)	0.551 (0.583)	210	0.325 (0.287)	0.279 (0.258)	210	0.356 (0.266)	0.263 (0.210)
18	0.326 (0.282)	0.352 (0.333)	231	0.316 (0.310)	0.289 (0.308)	0.547 (0.597)	207	0.345 (0.306)	0.293 (0.257)	207	0.372 (0.289)	0.271 (0.221)
21	0.338 (0.277)	0.367 (0.335)	227	0.328 (0.314)	0.300 (0.314)	0.544 (0.593)	203	0.367 (0.323)	0.301 (0.261)	203	0.381 (0.285)	0.271 (0.214)
24	0.353 (0.278)	0.377 (0.346)	224	0.344 (0.322)	0.318 (0.334)	0.544 (0.544)	200	0.393 (0.340)	0.302 (0.262)	200	0.403 (0.288)	0.279 (0.207)
36	0.365 (0.366)	0.436 (0.400)	122	0.388 (0.352)	0.380 (0.356)	0.727 (0.744)	117	0.489 (0.476)	0.320 (0.313)	117	0.494 (0.354)	0.279 (0.233)
48	0.448 (0.438)	0.494 (0.429)	102	0.477 (0.442)	0.416 (0.330)	0.771 (0.699)	99	0.553 (0.442)	0.377 (0.331)	99	0.521 (0.366)	0.304 (0.281)
59	0.526 (0.454)	0.488 (0.370)	93	0.544 (0.480)	0.469 (0.389)	0.735 (0.621)	91	0.554 (0.352)	0.397 (0.310)	91	0.510 (0.242)	0.331 (0.262)
60	0.519 (0.453)	0.477 (0.362)	89	0.539 (0.489)	0.477 (0.405)	0.745 (0.671)	86	0.551 (0.338)	0.401 (0.317)	86	0.507 (0.238)	0.336 (0.271)

Maturity (months)	2008 Oil Market Crash: Current Futures	2008: Average Futures	2008: Model	2008: Model q	2008: Schwartz 2F	2014 Oil Market Crash: Current Futures	2014: Average Futures	2014: Model	2014: Model q	2014: Schwartz 2F
1	0.275 (0.224)	0.642 (0.600)	0.254 (0.212)	0.272 (0.271)	0.619 (0.459)	0.187 (0.138)	0.338 (0.310)	0.182 (0.164)	0.268 (0.265)	0.336 (0.250)
3	0.56 (0.499)	0.773 (0.639)	0.509 (0.402)	0.508 (0.456)	1.777 (1.146)	0.363 (0.233)	0.503 (0.336)	0.394 (0.258)	0.468 (0.404)	0.966 (0.565)
6	0.797 (0.776)	0.807 (0.590)	0.677 (0.559)	0.658 (0.568)	4.029 (1.953)	0.481 (0.322)	0.712 (0.260)	0.544 (0.329)	0.440 (0.475)	2.351 (2.410)
9	0.57 (0.549)	0.576 (0.345)	0.463 (0.325)	0.414 (0.422)	5.940 (2.147)	0.583 (0.189)	0.861 (0.160)	0.688 (0.271)	0.383 (0.250)	4.735 (7.644)
12	0.428 (0.314)	0.437 (0.197)	0.346 (0.150)	0.286 (0.196)	8.123 (4.001)	0.707 (0.296)	1.024 (0.420)	0.806 (0.301)	0.430 (0.269)	8.578 (20.21)
15	0.372 (0.264)	0.369 (0.187)	0.295 (0.137)	0.293 (0.214)	11.68 (6.678)	0.752 (0.503)	1.101 (0.586)	0.901 (0.506)	0.447 (0.278)	10.83 (24.99)
18	0.337 (0.243)	0.360 (0.160)	0.290 (0.131)	0.276 (0.238)	17.01 (12.62)	0.691 (0.573)	1.080 (0.644)	0.871 (0.538)	0.395 (0.333)	16.16 (48.18)
21	0.352 (0.201)	0.390 (0.163)	0.291 (0.138)	0.227 (0.174)	24.04 (20.06)	0.521 (0.334)	0.822 (0.472)	0.650 (0.363)	0.276 (0.156)	20.13 (66.39)
24	0.403 (0.204)	0.390 (0.164)	0.335 (0.161)	0.322 (0.184)	31.76 (28.09)	0.615 (0.347)	0.930 (0.310)	0.679 (0.299)	0.267 (0.155)	29.84 (109.4)
36	0.314 (0.106)	0.296 (0.170)	0.306 (0.149)	0.263 (0.136)	123.2 (148.4)					
48	0.257 (0.147)	0.287 (0.114)	0.217 (0.115)	0.143 (0.150)	663.7 (1025)					
59	0.237 (0.128)	0.228 (0.153)	0.200 (0.140)	0.144 (0.145)	3579 (7229)					
60	0.224 (0.114)	0.226 (0.164)	0.204 (0.129)	0.156 (0.144)	4269 (9179)					

NOBS: Number of observations
Average Futures: Average of past 6 months futures
Model: Blind 2 factor model presented in this work
Model q: Forecast with known q

5.3 Case Analysis - Oil Market Crashes

This section of results is dedicated to analysis of model's performance during market crashes. Specifically, we focus on the 2008 and 2014 crashes, as identified in the figure 5.8 by vertical red lines. We analyze the model's forecasting performance on the day leading to the crash and average forecasting error over the next year after the crash. For this purpose, we change our calibration methodology in two ways; Firstly, we include all available information leading to crash by expanding our window analysis to include information from March 31st 1995 to the day of our analysis. Secondly, we move our window on a weekly rather than monthly basis to zoom in on the turbulent periods and analyze these periods in more detail. The expanding window calibration process is depicted in figure 5.9.

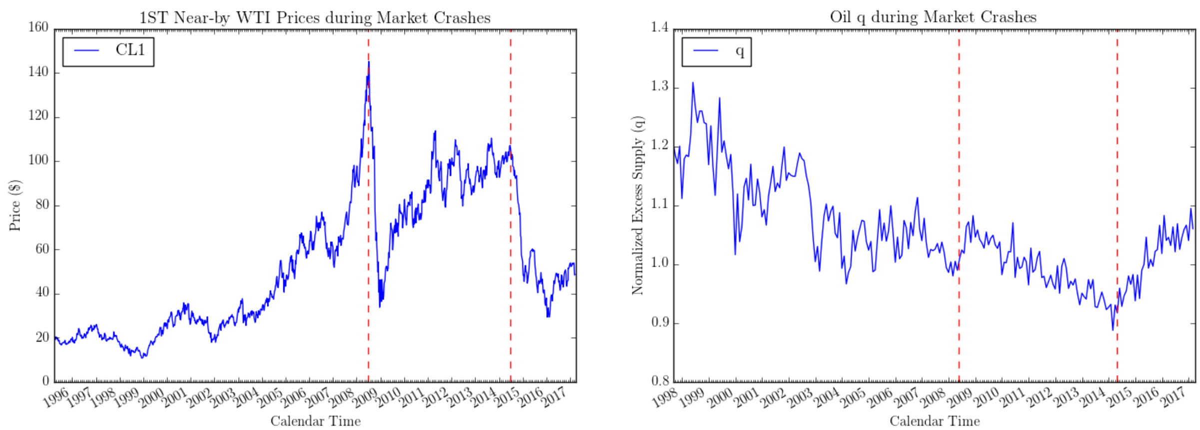


Figure 5.8: Oil Market Crashes of 2008 and 2014. Figures above demonstrate the prices and normalized excess supply for crude oil. The red lines mark the onset of the market crashes.

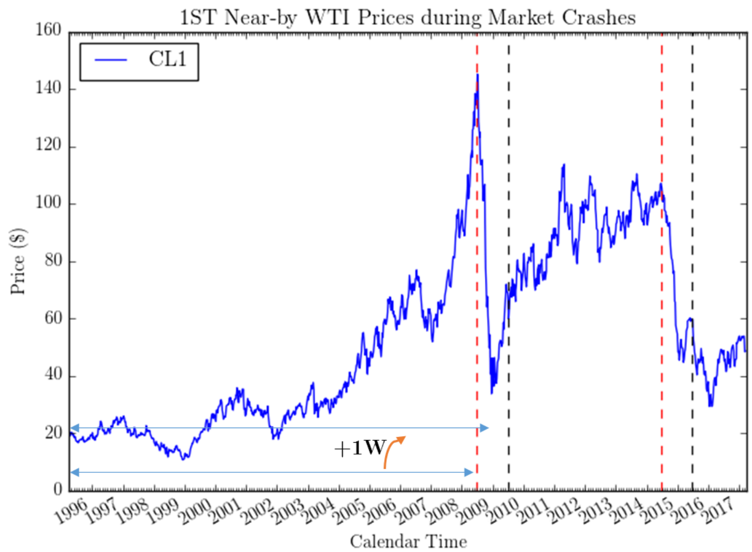


Figure 5.9: Expanding Window Example. The calibration window is moved one week at a time, beginning from the onset of the crash (red line) to the 1 year mark after crash (black line).

5.3.1 2008 Market Crash

The year 2008 witnessed massive swings in crude oil prices with prices reaching a peak of $145 and crashing to a low of $32. On July 15th 2008, the sell-off began and continued throughout the year. It is believed that executive order signed by the United State's president George W. Bush on July 15th lifting the ban on off-shore drilling initiated the sell-off. In the following months, easing of tensions between U.S. and Iran (July 2008) , Lehman Brothers bankruptcy (September 2008), a stronger U.S. Dollar (October 2008), decline in demand from the European Zone (October 2008), and increasing U.S. unemployment rate combined with weak economic data (November 2008) led to prices falling to $32 by the end of 2008.

5.3.2 2014 Market Crash

The year 2014 was filled with geopolitical tensions impacting oil prices. Commencement of the Joint Action Plan for Iran's nuclear program (January 2014), war in the middle east (Syria, Lybia, ...), and the dispute over Crimea and annexation of Crimea by Russia (February-March 2014) are among the headlines impacting oil prices in 2014. Finally in June 2014 the oil market boiled over and begun the recent crash. Increasing U.S. shale and light oil production, decreasing demand from China and the European Zone, and OPEC's inability to reach a deal over production cuts lead to oil prices crashing from $106 to $45 by the end of 2014.

5.3.3 Results During Market Crash Periods

Results of the out-of-sample forecasting for the two periods are presented in figure 5.10 and table 5.2. Figures 5.10a and 5.10c represent the actual out-of-sample forecasts on the day the crashes begun. For the 2008 crash, we can forecast up to 5 years while for the 2014 crash we could only extend the forecast horizon to two years. During both crashes, the model equipped with "q" information follows the prices closely, beating all other benchmarks. The blind model's performance is different during the two crashes. During the 2008 crash, the blind model forecasts the correct shape for prices and follows the trends while the forecasts are close to market expectations during the 2014 crash.

Furthermore, figures 5.10b and 5.10d present the average forecasts error for a 52-week period following the crash. Similarly, the results follow the same trends; The model equipped with "q" outperforms all the benchmarks while the blind model produces results with similar accuracy as the benchmark. In terms of standard error of forecasts, both the blind model and the model with q produce results with lower standard errors.

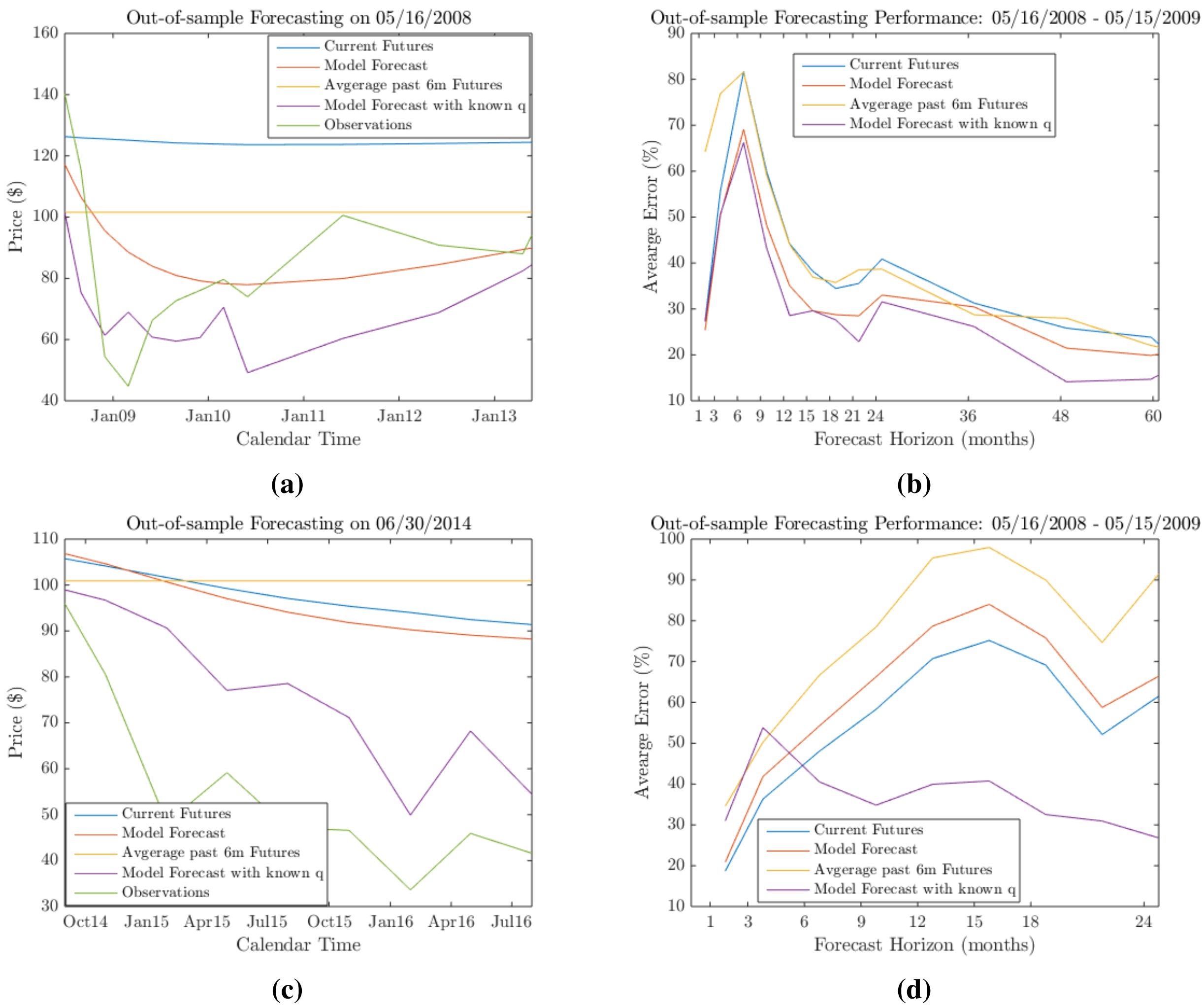


Figure 5.10: Out-of-sample Forecasting Performance During Market Crashes of 2008 and 2014. Figures on the left demonstrate the out-of-sample forecasts on the onset of the market crashes while the figures on the right represent the mean absolute percentage error during one year after the crash.

5.4 Scenario Analysis

In this section, we analyze the utility of our model for scenario analysis. As mention before, a unique contribution of the proposed model is the ability to analyze the impact of supply, demand, and inventory changes on values of derivatives such as futures contracts and options on futures contracts. Specifically, in light of ever-changing macroeconomic conditions, analysts can develop variant views of future consumption and production trends for the commodity. For instance, in case of Chinese economy's transition to consumerism, analysts can build scenarios for oil consumption trends and analyze the impact on commodity spot and derivative prices. Another example is the geopolitical tensions underlying oil production within OPEC countries. Deals regarding production quotas made under political tensions are fragile and not adhering to quotas could result in investment losses and increased risks for investors. In order to build out scenarios, assuming the current time period is denoted by "t", analysts can forecast variables in the following framework:

1. Forecast production for period $t+1$.

2. Forecast consumption for period $t+1$.

3. Adjust ending inventory for period $t+1$ as
$$Inventory_{t+1} = Inventory_t + Production_{t+1} - Consumption_{t+1}. \tag{5.1}$$

4. Forecast production and consumption for period $t+2$, taking into account producers' and consumers' reactions to the updated level of inventory and macroeconomic conditions.

5. Adjust ending inventory for period $t+2$, and repeat as necessary.

This framework is valuable to risk managers as well, providing a tool to hedge positions in the face of uncertain macroeconomic and political adversities. Here, we present an example of this framework applied to a series of positive shocks to oil production. Specifically, we develop two scenarios .

1. Scenario 1: Oil consumption and production continue their long-term trends assuming a linear model.

2. Scenario 2: We assume an initial breach in the recent OPEC production deal by a single member leading to increased production by all members over the next 12 months. The scenario is constructed as below.

   i Total production increases by $\frac{1}{12}$ million barrels per day, every month, for the next 12 months, as demonstrated in 5.11a.

   ii Consumption trend is assumed unchanged, as demonstrated in 5.11b.

   iii Inventory is built up periodically due to increased productionas demonstrated in 5.11c.

   iv An oil glut is made due to increased inventory and production. A new deal is reached bringing the production back to normal levels, a reduction of 1 million barrels per day by OPEC members.

   v Moreover, higher cost producers exit the market due to lower prices, reducing the production by another 0.5 million barrels per day.

Figure 5.11 demonstrates the two scenarios. Figure 5.11a compares the progress of production rate under the two scenarios, while figure 5.11b shows the steady consumption increases. Figures 5.11c and 5.11d show the inventory built up and impact on the normalized excess supply over the next two years.

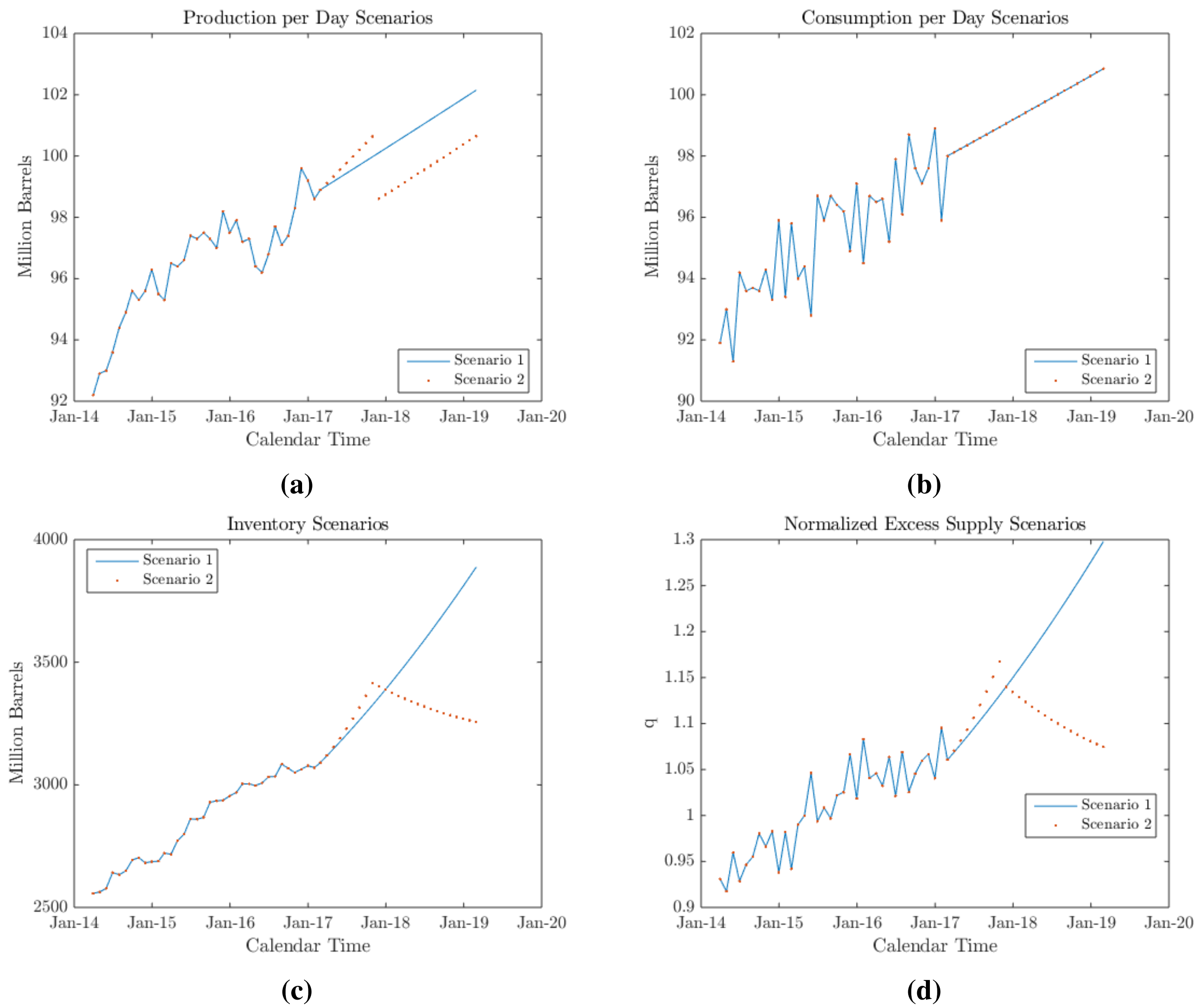


Figure 5.11: Scenario Analysis for increased production. The figures above present the production, consumption, inventory levels, and normalized excess supply under both scenarios. Values before January 2017 are historical values while the data after that is generated based on the scenarios.

Assuming the two scenarios, we can now forecast the spot prices. Figure 5.12 demonstrates the foretasted spot prices under both scenarios. Under the first scenario, production outpaced consumption, leading to significant price falls. Under the second scenario, extra production increases led to a sharper price decrease compared to the first scenario. However, after the first year, with the new OPEC deal and lower production of higher cost producers, total production falls by 1.5 million barrels per day. The production decrease leads to falling q and inventory levels while increasing prices. Furthermore, we can now estimate futures prices and options on futures. Figure 5.13 demonstrates the impact of scenarios on futures and options contracts. Specifically, we demonstrate the impact on constant maturity maturity futures contracts as well as the impact on at-the-money (ATM) call options on these futures contracts. The options are assumed to have the same maturity as the futures. Figures 5.13a and 5.13b demonstrate the impact on a short term 1 month maturity futures contract and the related ATM call option. Similarly, figures 5.13c and 5.13d demonstrate the impact on a 2 years maturity futures contract and the related ATM call option. Similar to spot prices, the futures prices and option prices fall under the first scenario, while the fall is sharper and is followed by a bounce back under the second scenario.

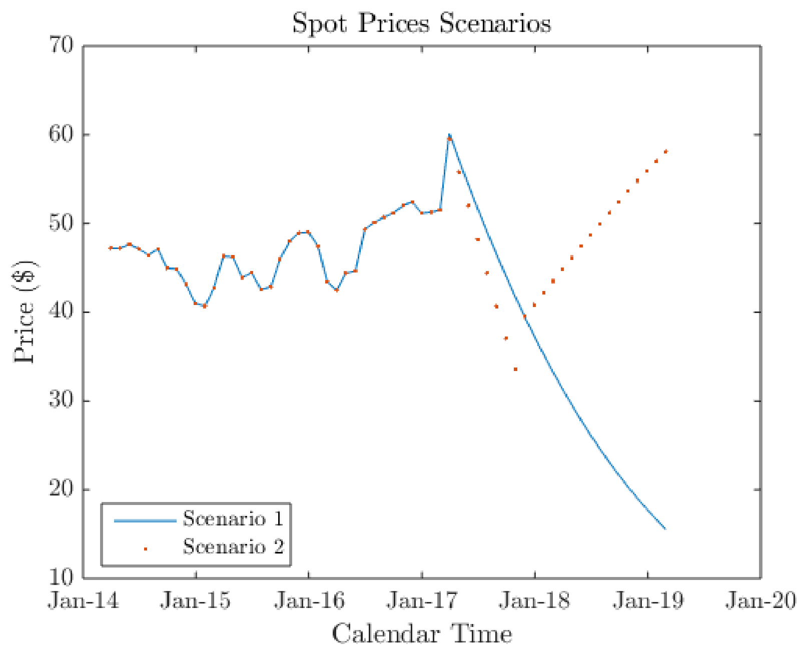


Figure 5.12: Scenario Analysis of Spot Prices. This figure presents the spot prices under the two proposed scenarios. Spot prices prior to January 2017 are results of calibration, combination of the two state variables in Kalman filter. Prices after January 2017 are produced under the two scenarios with model parameters calibrated to the latest period.

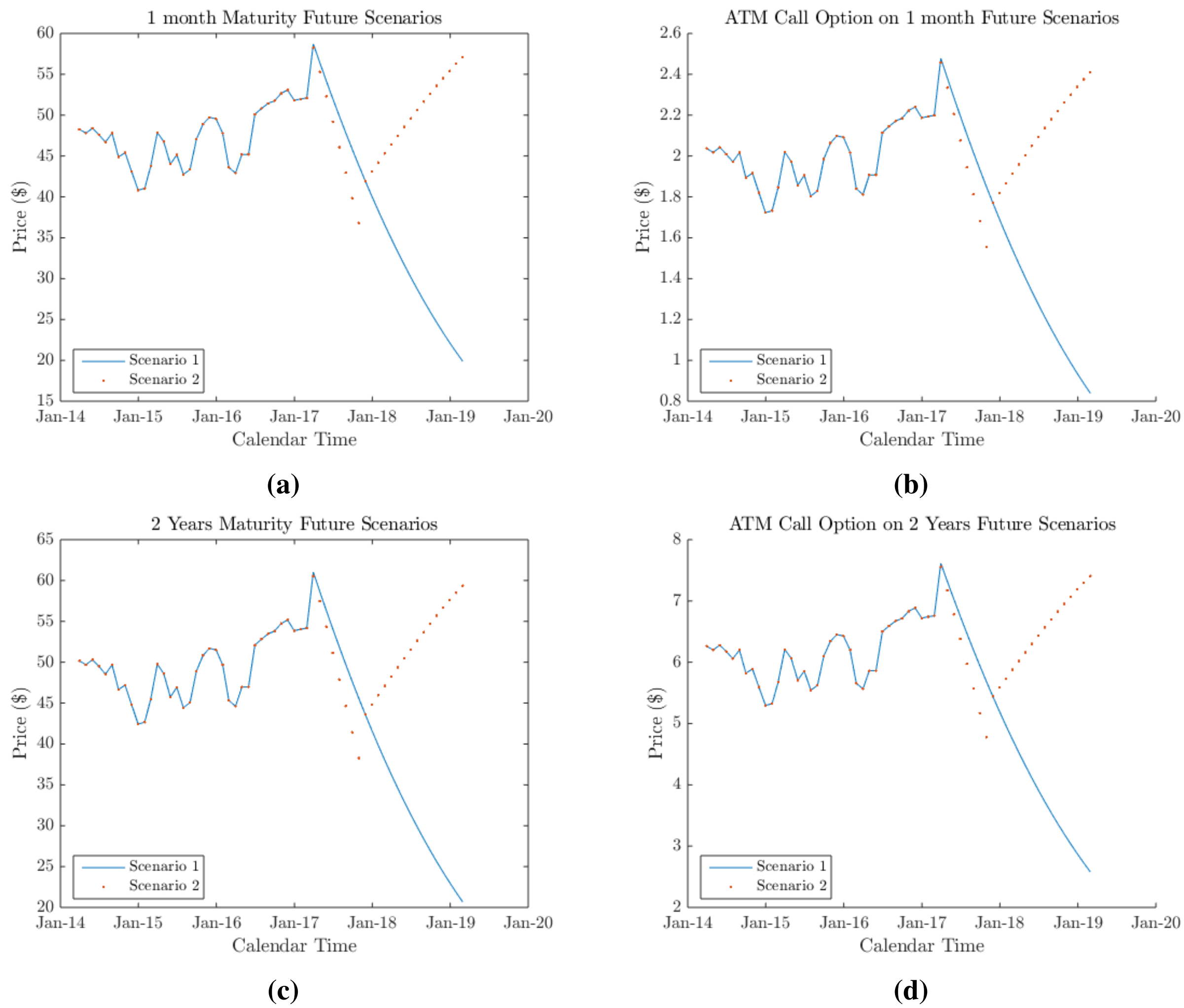


Figure 5.13: Scenario Analysis for futures and options. The figures above presents the implications of the two scenarios for futures and options on futures at 1-month and 2-years maturities.

6 Conclusions

In this paper, we put forward a framework to capture the observed dynamics between prices, supply/demand, and inventories. Specifically, normalized excess supply is proposed as the ratio of total excess supply to total demand. We show that normalized excess supply and log-prices follow an inverse linear relationship, consistent with economic theory of supply and demand. Furthermore, we show that normalized excess supply follows a mean reverting process. We propose that normalized excess supply determines the mean level of log-spot prices. In addition, deviations of prices from the mean level follow a mean reverting process. We utilize the aforementioned properties to develop a two factor stochastic model for estimation of storable commodity spot prices, with application in derivatives' valuations. The two stochastic factors of our model are normalized excess supply and the deviations from the mean level. Additionally, we develop two futures prices formulas based on the expectation theory as well as the no-arbitrage theory.

We calibrate the model to crude oil market data from March 1995 to March 2017 using a Kalman filter under three calibration setups; A 3-year rolling window, a 5-year rolling window, and an expanding window. We back test the model for both in-sample and out-of-sample data. We show that the proposed model matches the in-sample futures data very well and outperforms the market in out-of-sample foretasting of spot prices. Furthermore, we apply the model to the oil market crashes of 2008 and 2014, and demonstrate the models superiority in both accuracy and precision of out-of-sample forecasts. Additionally, we develop a scenario analysis framework to incorporate variant views of future production, consumption, and inventory levels into forecasting of spot prices. We further demonstrated the impact of the scenarios on futures and options on futures contracts. The spot, futures contracts, and option contracts prices are consistent with economic theory of supply and demand. We show the scenario analysis framework to be a valuable toolkit for investment management professionals and risk/quantitative analysts, allowing them to incorporate their views on macroeconomic developments in forecasting prices and valuations of investments.

References

Alquist, R., Kilian, L., and Vigfusson, R. J. (2013). Forecasting the Price of Oil. In *Handbook of Economic Forecasting*, pages 427–507.

Arismendi, J. C., Back, J., Prokopczuk, M., Paschke, R., and Rudolf, M. (2016). Seasonal stochastic volatility: Implications for the pricing of commodity options. *Journal of Banking & Finance*.

Bastian-Pinto, C., Brandão, L., and Hahn, W. J. (2009). Flexibility as a source of value in the production of alternative fuels : The ethanol case. *Energy Economics*, 31(3):411–422.

Baumeister, C. and Kilian, L. (2016). Real-Time Forecasts of the Real Price of Oil Real-Time Forecasts of the Real Price of Oil. *Journal of Business & Economic Statistics*, 30(2):326–336.

Beckmann, J., Belke, A., and Czudaj, R. (2014). Does global liquidity drive commodity prices? *Journal of Banking and Finance*, 48:224–234.

Belousova, J. and Dorfleitner, G. (2012). On the diversification benefits of commodities from the perspective of euro investors. *Journal of Banking and Finance*, 36(9):2455–2472.

Bera, A. K. and Jarque, C. M. (1987). A Test for Normality of Observations and Regression Residuals. *International Statistical Review*, 55(2):163–172.

BIS (2015). Semiannual OTC derivatives statistics. Technical report, Bank for International Settlements.

Black, F. (1976). The Pricing of Commodity Contracts. *Journal of Financial Economics*, 3:167–179.

Brandão, L. E. T., Penedo, G. M., and Bastian-Pinto, C. (2013). The value of switching inputs in a biodiesel production plant. *The European Journal of Finance, 2013*, 19(7-8):674–688.

Brennan, M. J. (1958). The Supply of Storage. *American Economic Review*, 48(1):50–72.

Bu, H. (2014). Effect of Inventory Announcements on Crude Oil Price Volatility. *Energy Economics*, 46:485–494.

Carlson, M., Khokher, Z., and Titman, S. (2007). Resource Price Dynamics. *The Journal of Finance*, LXII(4):1663–1703.

Casassus, J. and Collin-Dufresne, P. (2005). Stochastic Convenience Yield Implied from Commodity Futures and Interest Rates. *The Journal of Finance*, 60(5):2283–2331.

Casassus, J., Liu, P., and Tang, K. (2013). Economic linkages, relative scarcity, and commodity futures returns. *Review of Financial Studies*, 26(5):1324–1362.

Chen, S. and Insley, M. (2012). Regime switching in stochastic models of commodity prices: An application to an optimal tree harvesting problem. *Journal of Economic Dynamics and Control*, 36(2):201–219.

Chiang, I.-h. E., Hughen, W. K., and Sagi, J. S. (2015). Estimating Oil Risk Factors Using Information from Equity and Derivatives Markets. *The Journal of Finance*, LXX(2):769–804.

Cuñado, J. and de Gracia, F. P. (2003). Do oil price shocks matter? Evidence for some European countries. *Energy Economics*, 25(2):137–154.

Dickey, D. A. and Fuller, W. A. (1979). Distribution of the Estimators for Autoregressive Time Series With a Unit Root. *Journal of the American Statistical Association*, 74(366):427–431.

Ebert, L. and La Menza, T. (2015). Chile, copper and resource revenue: A holistic approach to assessing commodity dependence. *Resources Policy*, 43:101–111.

Elder, J., Miao, H., and Ramchander, S. (2012). Impact of macroeconomic news on metal futures. *Journal of Banking and Finance*, 36(1):51–65.

Erb, C. B. and Harvey, C. R. (2006). The Strategic and Tactical Value of Commodity Futures. *Financial Analysts Journal*, 62(2):69–97.

Geman, H. (2005). *Commodities and Commodity Derivatives: Modeling and Pricing for Agriculturals, Metals and Energy*. Wiley & Sons, Ltd., Chichester, West Sussex, England.

Geman, H. and Smith, W. O. (2013). Theory of storage, inventory and volatility in the LME base metals. *Resources Policy*, 38(1):18–28.

Gibson, R. and Schwartz, E. S. (1990). Stochastic Convenience Yield and the Pricing of Oil Contingent Claims.

Gorton, G. and Rouwenhorst, K. G. (2006). Facts and Fantasies about Commodity Futures. *Financial Analysts Journal*, 62(2):47–68.

Greer, R. J. (2000). The Nature of Commodity Index Returns. *THE JOURNAL OF ALTERNATIVE INVESTMENTS*, 3:45–53.

Güntner, J. H. F. (2014). How do oil producers respond to oil demand shocks ? *Energy Economics*, 44:1–13.

Hamilton, J. D. (2008). Oil and the Macroeconomy. In Durlauf, S. N. and Blume, L. E., editors, *In The New Palgrave Dictionary of Economics*. Palgrave Macmillan, 2 edition.

Hamilton, J. D. (2011). Nonlinearities and the Macroeconomic Effects of Oil Prices. *Macroeconomic Dynamics*, 15(Supplement 3):364–378.

Hikspoors, S. and Jaimungal, S. (2008). Asymptotic Pricing of Commodity Derivatives using Stochastic Volatility Spot Models. *Applied Mathematical Finance*, 15(5-6):449–477.

Hong, H. and Yogo, M. (2012). What does futures market interest tell us about the macroeconomy and asset prices? *Journal of Financial Economics*, 105(3):473–490.

Hurst, H. E. (1951). Long-term storage capacity of reservoirs. *Transactions of American Society of Civil Engineers*, 116:770–808.

Irwin, S. H. and Sanders, D. R. (2011). Index funds, financialization, and commodity futures markets. *Applied Economic Perspectives and Policy*, 33(1):1–31.

Juvenal, L. and Petrella, I. (2015). Speculation in the Oil Market. *Journal of Applied Econometrics*, 30(4):621–649.

Kaminski, V. (2014). The microstructure of the North American oil market. *Energy Economics*, 46:S1–S10.

Kilian, L. (2014). Oil Price Shocks : Causes and Consequences. *The Annual Review of Resource Economics*, 6:133–154.

Kilian, L. and Lee, T. K. (2014). Quantifying the speculative component in the real price of oil: The role of global oil inventories. *Journal of International Money and Finance*, 42:71–87.

Kilian, L. and Murphy, D. P. (2014). The Role of Inventories and Speculative Trading in the Global Market for Crude Oil. *Journal of Applied Econometrics*, 29(3):454–478.

Kobari, L., Jaimungal, S., and Lawryshyn, Y. (2014). A real options model to evaluate the effect of environmental policies on the oil sands rate of expansion. *Energy Economics*, 45:155–165.

Lai, A. N. and Mellios, C. (2016). Valuation of commodity derivatives with an unobservable convenience yield. *Computers and Operations Research*, 66:402–414.

Liu, P. and Tang, K. (2011). The stochastic behavior of commodity prices with heteroskedasticity in the convenience yield. *Journal of Empirical Finance*, 18(2):211–224.

Mandelbrot, B. B. (2003). Heavy Tails in Finance for Independent or Multifractal Price Increments. In S.T. Rachev, editor, *Handbook of Heavy Tailed Distributions in Finance*, chapter 1. Elsevier Science B.V.

Mirantes, A. G., Población, J., and Serna, G. (2013). The stochastic seasonal behavior of energy commodity convenience yields. *Energy Economics*, 40:155–166.

Morana, C. (2013). The Oil Price-Macroeconomy Relationship Since the Mid-1980s: A Global Perspective. *The Energy Journal*, 34(3):153–189.

Paschke, R. and Prokopczuk, M. (2010). Commodity derivatives valuation with autoregressive and moving average components in the price dynamics. *Journal of Banking and Finance*, 34(11):2742–2752.

Ratti, R. A. and Vespignani, J. L. (2015). Commodity prices and BRIC and G3 liquidity: A SFAVEC approach. *Journal of Banking and Finance*, 53:18–33.

Routledge, B. R., Seppi, D. J., and Spatt, C. S. (2000). Equilibrium Forward Curves for Commodities. *The Journal of Finance*, LV(3):1297–1338.

Schwartz, E. S. (1997). The Stochastic Behavior of Commodity Prices: Implications for Valuation and Hedging. *The journal of Finance*, 52(3):923 – 973.

Sidebottom, A., Belur, J., Bowers, K., Tompson, L., and Johnson, S. D. (2011). Theft in Price-Volatile Markets: On the Relationship between Copper Price and Copper Theft. *Journal of Research in Crime and Delinquency*, 48(3):396–418.

Sockin, M. and Xiong, W. (2015). Informational Frictions and Commodity Markets. *The Journal of Finance*, LXX(5):2063–2098.

Stefanski, R. (2014). Structural transformation and the oil price. *Review of Economic Dynamics*, 17(3):484–504.

Szymanowska, M., Roon, F. D. E., Nijman, T., and Boorbergh, R. V. D. (2014). An Anatomy of Commodity Futures Risk Premia. *The Journal of Finance*, LXIX(1):453–482.

Telser, L. G. (1958). Futures Trading and the Storage of Cotton and Wheat. *Journal of Political Economy*, 66(3):233–255.

Tilton, J. E., Humphreys, D., and Radetzki, M. (2012). Investor demand and spot commodity prices: Reply. *Resources Policy*, 37(3):397–399.

Trolle, A. B. and Schwartz, E. S. (2009). Unspanned Stochastic Volatility and the Pricing of Commodity Derivatives. *The Review of Financial Studies*, 22(11):4423–4461.

7 Appendix

7.1 Derivation of State Distributions

This appendix represents the derivation for join distribution of normalized excess supply and deviations from the mean level. The state variables follow a multivariate normal distribution.

Joint stochastic process for normalized excess supply and log-price expressed in equations 3.2 and 3.3 can be decomposed as

$$dq = \kappa_1(\theta_1 - q)dt + \sigma_1 dz_1, \tag{7.1}$$

$$d\varepsilon = -\kappa_2 \varepsilon dt + \sigma_2 \rho dz_1 + \sigma_2 \sqrt{1-\rho^2} dz_2, \tag{7.2}$$

with independent standard Brownian motions z_1 and z_2.

Solution to normalized excess supply can be obtained as a general Ornstein-Uhlenbeck processes

$$q_t = \mathrm{e}^{-\kappa_1 t} q_0 + \theta_1 \left(1 - \mathrm{e}^{-\kappa_1 t}\right) + \sigma_1 \mathrm{e}^{-\kappa_1 t} \int_0^t \mathrm{e}^{\kappa_1 u} dz_1, \tag{7.3}$$

Solution to ε_t could be obtained as

$$\varepsilon_t = \mathrm{e}^{-\kappa_2 t} \varepsilon_0 + \sigma_2 \rho \mathrm{e}^{-\kappa_2 t} \int_0^t \mathrm{e}^{\kappa_2 s} dz_1 + \sigma_2 \sqrt{1-\rho^2} \mathrm{e}^{-\kappa_2 t} \int_0^t \mathrm{e}^{\kappa_2 s} dz_2 \tag{7.4}$$

Replacing q_t and ε_t in equation 3.1 with the solutions above, we obtain

$$\begin{aligned} X_t &= a\left(\left(1-\mathrm{e}^{-t\kappa_1}\right)\theta_1+\mathrm{e}^{-t\kappa_1}q_0\right)+\mathrm{e}^{-t\kappa_2}\varepsilon_0+b \\ &+\int_0^t\left(a\sigma_1\mathrm{e}^{-\kappa_1(t-s)}+\sigma_2\rho\mathrm{e}^{-\kappa_2(t-s)}\right)dz_1 \\ &+\sigma_2\sqrt{1-\rho^2}\mathrm{e}^{-\kappa_2 t}\int_0^t\mathrm{e}^{\kappa_2 s}dz_2 \end{aligned} \tag{7.5}$$

The following moments are then obtained for q, ε, and X

$$\mu_q=\mathbf{E}\left[q_t\right]=\mathrm{e}^{-\kappa_1 t}q_0+\theta_1\left(1-\mathrm{e}^{-\kappa_1 t}\right), \tag{7.6}$$

$$\mu_\varepsilon=\mathbf{E}\left[\varepsilon_t\right]=\mathrm{e}^{-\kappa_2 t}\varepsilon_0 \tag{7.7}$$

$$\mu_X=\mathbf{E}\left[X_t\right]=a\left(\mathrm{e}^{-\kappa_1 t}q_0+\theta_1\left(1-\mathrm{e}^{-\kappa_1 t}\right)\right)+\mathrm{e}^{-\kappa_2 t}\varepsilon_0+b \tag{7.8}$$

$$\sigma_q^2=\mathbf{Var}[q_t]=\frac{1}{2}\frac{\sigma_1^2\left(1-\mathrm{e}^{-2\kappa_1 t}\right)}{\kappa_1}, \tag{7.9}$$

$$\sigma_\varepsilon^2=\mathbf{Var}[\varepsilon_t]=\frac{1}{2}\frac{\sigma_2^2\left(1-\mathrm{e}^{-2\kappa_2 t}\right)}{\kappa_2}, \tag{7.10}$$

$$\sigma_X^2=\frac{a^2{\sigma_1}^2\left(1-\mathrm{e}^{-2t\kappa_1}\right)}{2\kappa_1}+\frac{{\sigma_2}^2\left(1-\mathrm{e}^{-2t\kappa_2}\right)}{2\kappa_2}+2\frac{a\sigma_1\,\sigma_2\,\rho\left(1-\mathrm{e}^{-t(\kappa_1+\kappa_2)}\right)}{\kappa_1+\kappa_2} \tag{7.11}$$

$$\sigma_{q\varepsilon}=\mathbf{Cov}[q_t,\varepsilon_t]=-\frac{\sigma_1\,\sigma_2\,\rho\left(\mathrm{e}^{-t(\kappa_1+\kappa_2)}-1\right)}{\kappa_1+\kappa_2} \tag{7.12}$$

Under the risk-neutral measure, only μ_q and μ_X require to be updated. Under this measure

$$\mu_q^Q=\mathbf{E}^Q\left[q_t\right]=\mathrm{e}^{-\kappa_1 t}q_0+\hat{\theta}_1\left(1-\mathrm{e}^{-\kappa_1 t}\right), \tag{7.13}$$

$$\mu_\varepsilon^Q=\mathbf{E}^Q\left[\varepsilon_t\right]=\mathrm{e}^{-t\kappa_2}\varepsilon_0-\frac{\lambda_2}{\kappa_2}\left(1-\mathrm{e}^{-t\kappa_2}\right), \tag{7.14}$$

$$\mu_X^Q=\mathbf{E}^Q\left[X_t\right]=a\left(\left(1-\mathrm{e}^{-t\kappa_1}\right)\hat{\theta}_1+\mathrm{e}^{-t\kappa_1}q_0\right)-\frac{\lambda_2}{\kappa_2}\left(1-\mathrm{e}^{-t\kappa_2}\right)+\mathrm{e}^{-t\kappa_2}\varepsilon_0+b, \tag{7.15}$$

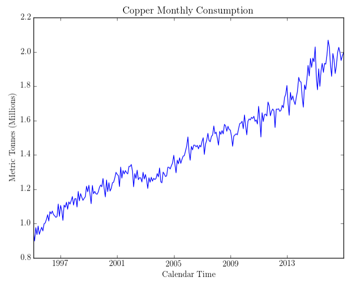
where

$$\hat{\theta}_1=\theta_1-\frac{\lambda_1}{\kappa_1}, \tag{7.16}$$

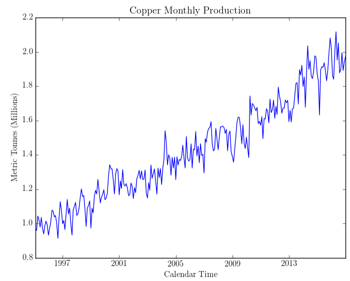
7.2 Derivation of Futures Options Prices

7.3 Industrial Metals Consumption, Production, and Inventory

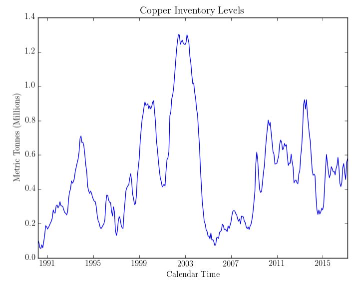
This appendix presents the global consumption, production, and inventory for industrial metals. Inventory data is collected from major metal exchanges, the London Metal Exchange (LME), the Commodity Exchange (COMEX), and the Shanghai Futures Exchanges (SHFE).



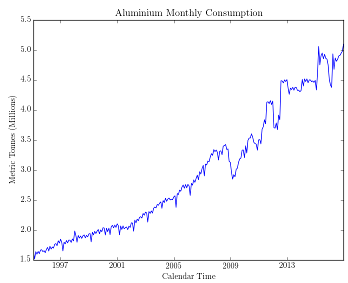
(a)



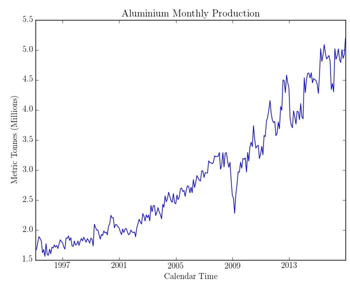
(b)



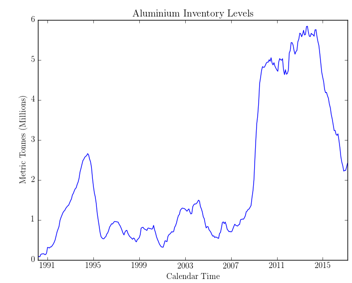
(c)



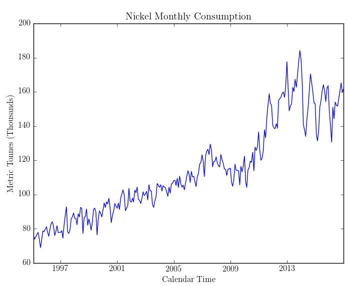
(d)



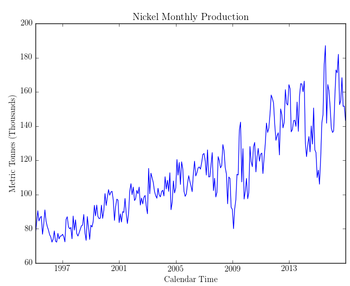
(e)



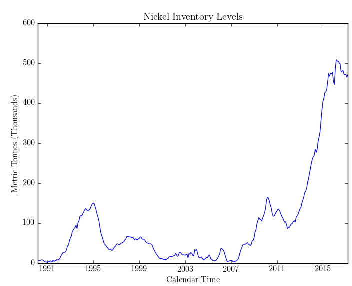
(f)



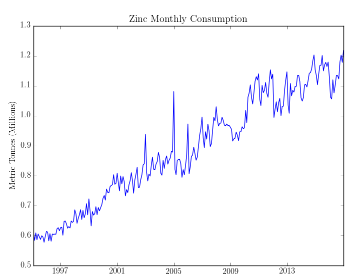
(g)



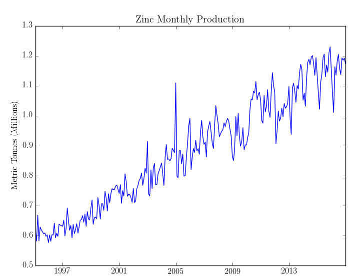
(h)



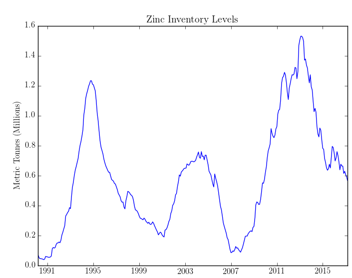
(i)



(j)



(k)



(l)

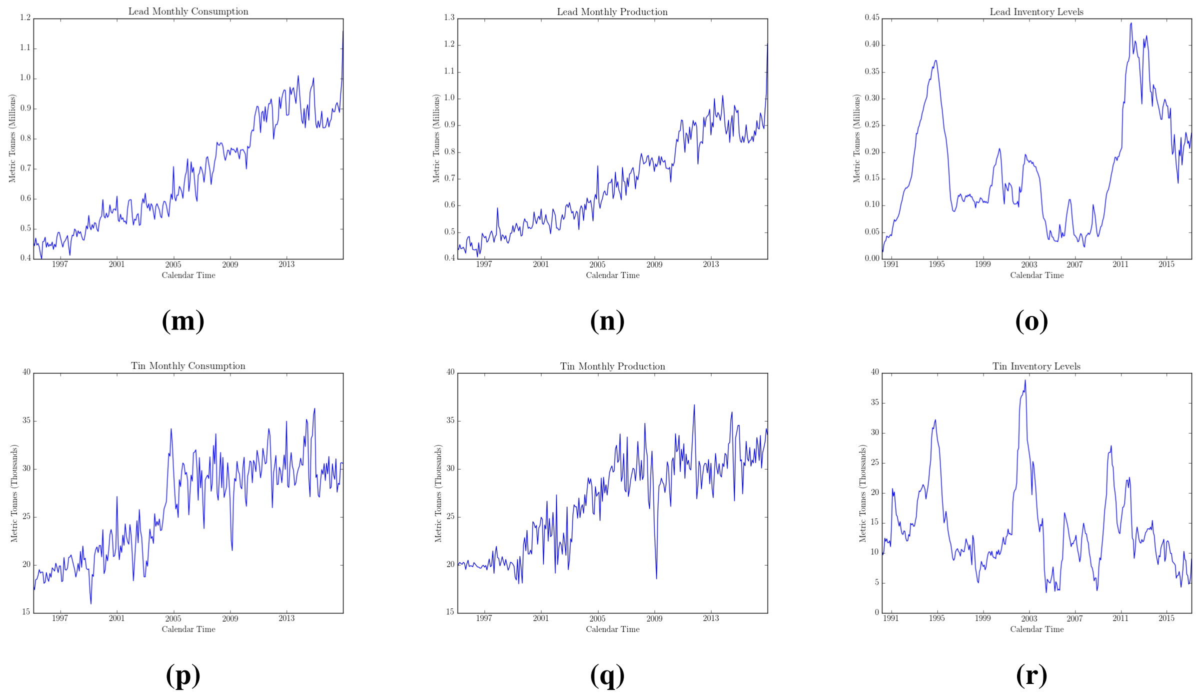
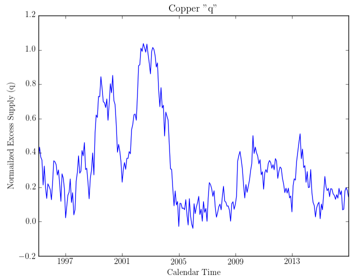


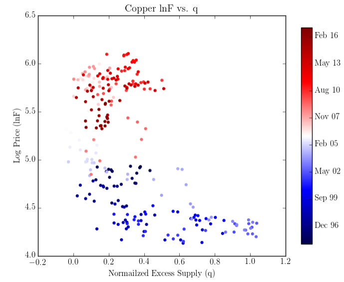
Figure 7.1: Plots of Consumption, Production, and Total Inventory for Industrial Metals. The total inventory is calculated as sum of available inventory at LME, COMEX, and Shanghai Exchange.

7.4 Industrial Metals Normalized Excess Supply and Deviation from Mean

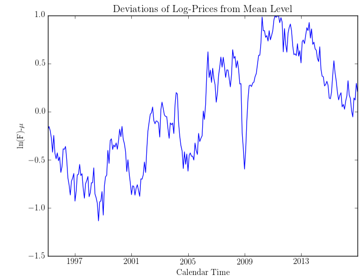
This appendix presents the normalized excess supply (q), the relationship between prices and "q", and the deviations from the mean level for industrial metals.



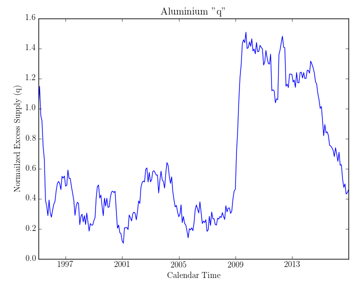
(a)



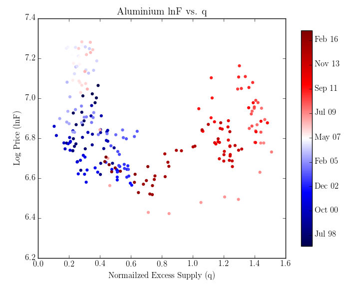
(b)



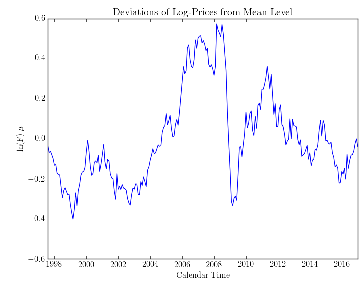
(c)



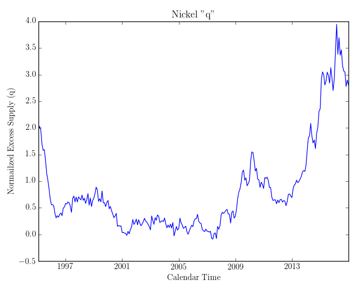
(d)



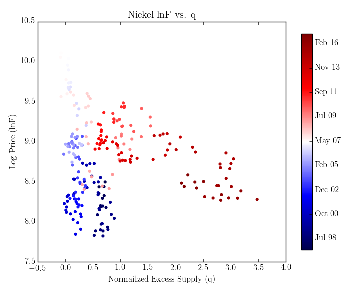
(e)



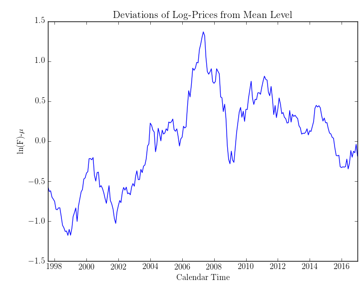
(f)



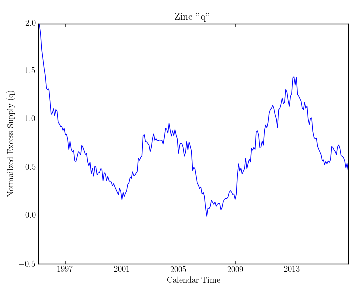
(g)



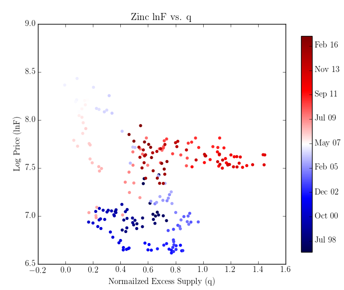
(h)



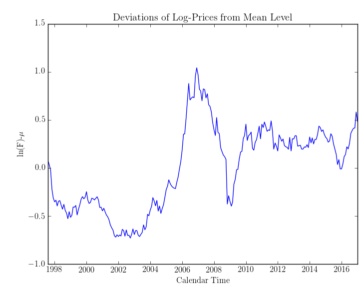
(i)



(j)



(k)



(l)

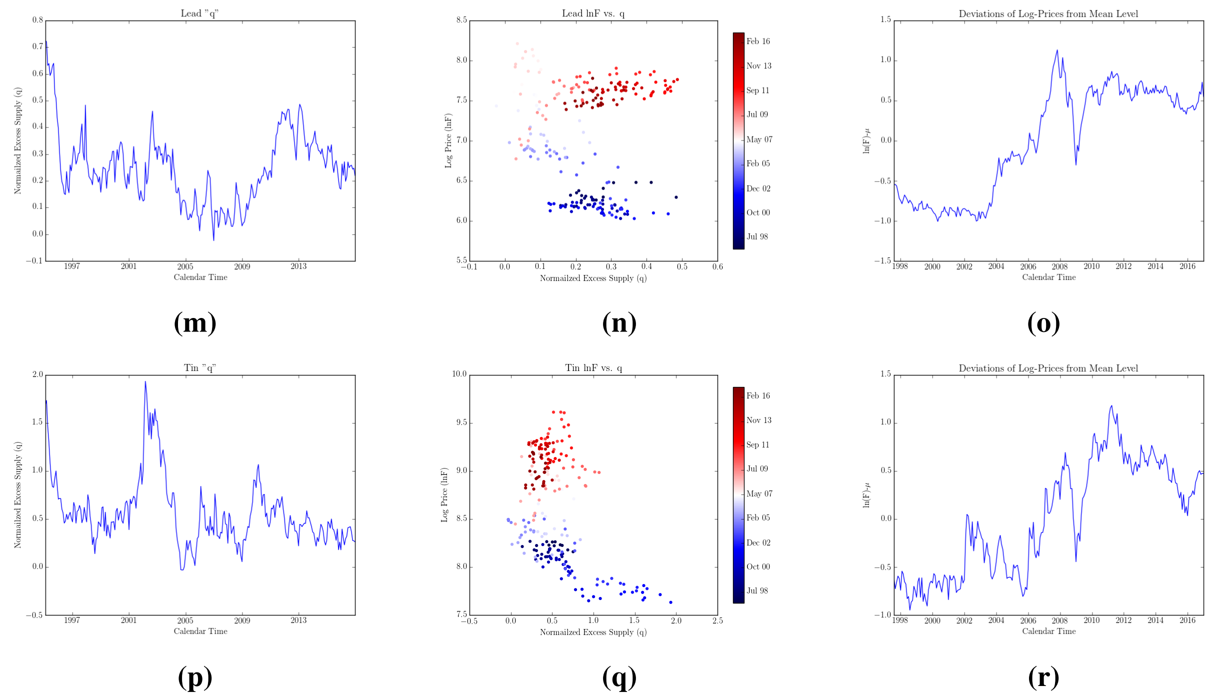


Figure 7.2: Industrial Metals q, $\ln(F) vs. q$ and Price Deviations from Mean Level. Figures above demonstrate the monthly normalized excess supply, first nearby future vs. "q", and ε for industrial metals.

7.5 Calibration

For the purpose of calibration, a Kalman filter process similar to that of Schwartz (1997) is utilized. The two underlying stochastic factors in the proposed model are normalized excess supply, q, and the deviation from mean level of log-prices, ε. Spot prices are unobservable for oil. Often, first nearby futures' prices are used as a proxy for the spot prices. Moreover, the published supply, demand, and inventory data are only estimates of the true unobservable variables. This makes straightforward calibration of the processes inaccurate. Hence, the model is put in a state-space form to account for unknown true values of the state variables. The two state variables are the q and ε. The Kalman filter is then applied to estimate the true value of the state variables' time series. The Kalman filter is an iterative prediction-correction algorithm. In each time step t, based on the current estimate of parameters and state variables, the time $t+1$ value of state variables are estimated through a transition equation. The measurement variables are then estimated based on the predicted state variables through the measurement equation. Finally, the predicted state variables are corrected based on the prediction error, the difference between observed and predicted data. The transition equation relating the state parameters at time t to the state parameters at time $t-\Delta t$ is set up as

$$\bar{x}_t^- = A_t \bar{x}_{t-\Delta t} + B_t + w_t \tag{7.17}$$

where

$$\bar{x}_t^- = \begin{bmatrix} q_t^- \\ \varepsilon_t^- \end{bmatrix}, \tag{7.18}$$

$$A_t = \begin{bmatrix} \mathrm{e}^{-\kappa_1 t} & 0 \\ 0 & \mathrm{e}^{-\kappa_2 t} \end{bmatrix}, \tag{7.19}$$

$$B_t = \begin{bmatrix} \theta_1 (1-\mathrm{e}^{-\kappa_1 t}) \\ 0 \end{bmatrix}, \tag{7.20}$$

and w_t is defined as transition noise with

$$E[w_t] = 0_{2\times 2}, Var[w_t] = \begin{bmatrix} \sigma_q^2 & \sigma_{q\varepsilon} \\ \sigma_{\varepsilon q} & \sigma_\varepsilon^2 \end{bmatrix}. \tag{7.21}$$

Subsequently, the measurement equation relating the observables to state variables is set up as

$$\bar{y}_t = C_\tau \bar{x}_t^- + D_t + v_t, \tag{7.22}$$

where

$$\bar{y}_t = \begin{bmatrix} \ln\left(F\left(t, \tau_i\right)\right) \\ q_t \end{bmatrix}, \tag{7.23}$$

$$C_\tau = \begin{bmatrix} a & 1 \\ 1 & 0 \end{bmatrix}, \tag{7.24}$$

$$D_\tau = \begin{bmatrix} \left(r_{\tau_i} - \delta_{\tau_i}\right)\tau_i \\ 0 \end{bmatrix}, \tag{7.25}$$

and v_t is defined as measurement noise with

$$E\left[v_t\right] = 0_{(n+1)\times(n+1)}, Var\left[v_t\right] = \begin{bmatrix} R_1 & & \\ & \ddots & \\ & & R_{n+1} \end{bmatrix}, \tag{7.26}$$

where R_i is measurement noise associated with i^{th} observation.

Equations 7.17 to 7.26 set out the prediction stage of the Kalman filter. The correction stage involves the following equations

$$P_t^- = A_t P_{t-1} A_t' + Q_{t-1}, \tag{7.27}$$

$$G_t = P_t^- C_t' \left(C_t P_t^- C_t' + v_t\right), \tag{7.28}$$

$$\bar{x}_t = \bar{x}_t^- + G_t\left(\bar{y}_t - D_t - C_t \bar{x}_t^-\right), \tag{7.29}$$

and

$$P_t = \left(I - G_t C_t\right) P_t^-. \tag{7.30}$$

Finally, this iterative system is passed to an optimization engine to maximize the likelihood function over the underlying parameters. The negative log-likelihood function is then derived as

$$\frac{n+1}{2}T\ln(2\pi)+\frac{1}{2}\sum_{all\ t}\ln(\det(V_t))+e_t^{'}V_t^{-1}e_t, \tag{7.31}$$

where

$$V_t = C_t P_t^{-} C_t^{'} + v_t, \tag{7.32}$$

and

$$e_t = \bar{y}_t - D_t - C_t \bar{x}_t^{-}. \tag{7.33}$$

Lastly, the maximum likelihood estimate of the parameters, the standard error of the parameters and the optimal estimates of the state variables are obtained.