An Optimal Solution to a General Dynamic Jet Fuel Hedging Problem

Juliana M. Nascimento
Warren B. Powell
Department of Operations Research and Financial Engineering
Princeton University

February 11, 2008
Abstract

We propose a dynamic hedging strategy for jet fuel which strikes a balance between hedging against jumps in the price of jet fuel and placing bets that the price will rise, lowering the overall cost of jet fuel. We model the commodity price using an unobservable two-factor model that allows mean-reversion in short-term prices and uncertainty in the equilibrium level to which prices revert. We combine dynamic programming and Kalman filter estimation to obtain an optimal policy that minimizes the expected costs while keeping the variance at low levels.
1 Introduction

Jet fuel costs account for a large portion of an airline’s operating expenses and when fuel prices rise dramatically, airlines cannot pass all of the costs on to their customers (Zea (2004)). Also, it is almost impossible for an airline to stock large amounts of jet fuel, due to financing and storage costs. Therefore, an effective strategy for airlines is to hedge fuel costs to avoid huge swings in expenses. As pointed out in Morrell & Swan (2006), hedging may still help reduce volatility in earnings by moving profits from one quarter to another.

Despite its potential, jet fuel hedging remains a largely unused strategy. Some executives claimed that risk would be present regardless of whether they hedged or not. Furthermore, they claimed that hedging was not a core competency and as long as competitors were not hedged, it was a level playing field (Zea (2004)). But low-cost airlines such as Southwest (Carter et al (2004)) have benefited considerably from an aggressive hedging strategy. It was shown in Carter et al. (2006) that there is a hedging premium for stocks of airlines using derivatives to hedge jet fuel exposure. The authors also show that hedging allows airlines to take advantage of investment opportunities in times of high commodity prices.

Both over-the-counter (OTC) and exchange-traded derivatives can be used for hedging. Nevertheless, a perfect hedge is practically impossible. OTC contracts on jet fuel include options, collar structures and swaps. Even though the ability to customize these derivatives is a big plus, there are several disadvantages associated with them. First, they are rather illiquid, making them very expensive. Second, there is counter-party risk. Finally, they are not available in quantities sufficient to hedge all of an airline’s jet fuel consumption.

On the other hand, exchange-traded derivatives are more liquid and eliminate counter-party risk. However, these contracts are not available in the U.S. for jet fuel. Therefore, contracts on commodities that have a high price correlation with jet fuel must be used for hedging. Heating and crude oil are usually the commodities of choice, since jet fuel shares similar characteristics with the first and is refined from the second. Finally, as exchange-traded contracts for hedging jet fuel costs are standardized (inflexible) and are based on

\footnote{See the article An Airline Shrugs at Oil Prices at The New York Times on November 29, 2007.}
a different underlying commodity, the associated basis risk is larger compared with OTC derivatives.

In this paper, we derive a dynamic strategy based on exchange-traded futures contracts on heating or crude oil to hedge jet fuel demand that will occur at time $T$. The hedging policy should maximize

$$
E \left[ \sum_{t \in T} e^{-rt} U_t(W_t, R_t, X_t) \right],
$$

where $T$ is the set of trading times prior to $T$, $r$ is the risk-free interest rate, $W_t$ is the information available up until time $t$, $R_t$ is the number of futures contracts standing at $t$ and $X_t$ is the policy. The information vector $W_t$ includes the spot price of oil and fuel. It also includes the current and the previous (last trading date) price of the future contract maturing a month after $T$. The dynamic policy $X_t$ determines the trading decision taken at $t$. Moreover, $U_t(W_t, R_t, X_t)$ is the utility function under consideration, where both risk and return play a role. A prospective test on the hedge effectiveness is also taken into account. The relative importance of risk versus the cost of the jet fuel is defined by a weight determined by the investor to accommodate his preferences.

The most widely used hedging strategy considering futures contracts is the dynamic minimum variance optimal hedge ratio ($OHR_t$) given by

$$
OHR_t = \frac{\sigma_{tFC}}{\sigma_{tF}^2},
$$

where, in the jet fuel case, $\sigma_{tFC}$ is the covariance between the price of the future contract and the spot price of fuel and $\sigma_{tF}^2$ is the variance of the price of the future contract. Both the covariance and the variance are computed using information available up until time $t$. Almost all the research on hedging focuses on minimizing the variance of the deviation between jet fuel and the hedge (Lien & Tse (2002)), ignoring other dimensions such as the cost of the jet fuel.

We differ from this conventional approach since we focus on deriving and computing an optimal dynamic hedging strategy for a more general utility which considers not only the variance of the deviation between jet fuel and the hedge, but also the expected cost of the
jet fuel. Thus, depending on the weight that we put on the variance, we are willing to accept higher risk if it lowers the expected cost of the jet fuel. We use dynamic programming to derive the optimal strategy.

To model the oil spot price we use the two-factor model proposed in [Schwartz & Smith (2000)]. This model allows mean-reversion in short-term prices and uncertainty in the equilibrium level to which prices revert. To model the relationship between the oil spot price and the fuel spot price, we use the time varying regression (TVR) model discussed in [Bos & Gould (2007)]. In this model, the correlation between the two prices varies with time and is represented by a martingale. However, the two factors in the oil spot price model can not be directly observed and their parameters are not known. The same is true for the martingale and the associated parameters. We use the Kalman filter and maximum likelihood estimation to recursively estimate the parameters and the processes themselves based on observations of spot and future prices for oil and spot prices for fuel.

The main contribution of this paper is the derivation of an optimal dynamic hedging strategy with respect to a utility function that allows the investor to control the relative importance of risk mitigation and speculation. Even though the curse of dimensionality of dynamic programming could have prevented us from actually computing such policy, we show that the corresponding optimal value function is quadratic in the number of contracts standing. We also show that computing the optimal policy depends only on our ability to compute the expected values involving the future contract prices and the spot prices of fuel and oil. The two-factor [Schwartz & Smith (2000)] and the TVR [Bos & Gould (2007)] models allow us to obtain accurate estimates of the expected values. Using historical price data, we demonstrate the relative performance of our policy over popular competing policies as a function of the weight on risk. For certain weight ranges, we show that our policy dominates other policies by producing hedges with lower risk as well as lower average fuel cost.
2 Literature Review on Hedging

There is an extensive literature on the use of derivatives to hedge a particular risk in the future ([Lien & Tse (2002)]). The risk might be associated to a foreign exchange rate, to a commodity price or to the level of the stock market to name a few. Jet fuel hedging is a special case, with its own characteristics. The best-known hedging strategy using futures contracts ([Hull (2000)]) is the minimum variance optimal hedge ratio (OHR), which is the ratio of the size of the position taken in future contracts to the size of the exposure. The OHR is usually used as a benchmark. In [Cecchetti et al. (1988)] the authors applied the two-asset portfolio model to determine that

\[
OHR = \frac{\sigma_{FC}}{\sigma_F^2},
\]

where, as described in the introduction, \(\sigma_{FC}\) is the covariance between the price of the future contract and the spot price of fuel and \(\sigma_F^2\) is the variance of the price of the future contract. This is a static strategy called *hedge-and-forget*. It assumes that the hedge position is not going to be changed over the duration of the hedge. The choice of the model and the technique to estimate the covariance and the variance is the subject of several papers. For example, in [Tunaru & Tan (2002)], the authors use different regression models.

On the other hand, a large literature initiated by [Baillie & Myers (1991)] allows the minimum variance hedge ratio to vary over time. The dynamic version is given by (2). A short review on estimating the time-dependent (co)variances can be found in [Chen et al. (2003)]. Most papers estimate dynamic hedge ratios ([Bos & Gould (2007)]) using either the (G)ARCH framework of [Engle (1982) and Bollerslev (1986)] or the stochastic volatility approach, introduced into the econometric literature by [Harvey et al. (1994) and Jacquier et al. (1994)]. The required estimations are obtained either using a (quasi) maximum likelihood or a Bayesian approach.

Combining the GARCH and the dynamic programming frameworks to construct a dynamic strategy has been explored before. In [Haigh & Holt (2002)], the authors combined the two to hedge the risk associated with purchasing commodities for use in food manufacturing. Their objective was to minimize the variance of the total terminal cost. The resulting
optimal policy is very similar to the dynamic OHR. The only difference, for time period $t$, is that the variance in the denominator of (2) is multiplied by $r^{T-t}$, where $r$ is a discount factor and $T$ is the end of the hedge period. Our paper differs from this work since we consider a more complex utility function, yielding a strategy that allows for a tradeoff between hedging and speculation according to the investor’s preferences.

Both the static and the dynamic hedge ratios are optimum when the objective is to minimize the risk, not to maximize expected utility, which also depends on expected return. As pointed out by Cecchetti et al. (1988), only a totally risk averse investor can make an optimal hedging decision without taking the impact on both risk and return into account. Value at risk (VaR) is another metric that has been considered in the literature. The static and dynamic zero-VaR hedge ratios for futures contracts were discussed in Hung et al. (2005) and Lee & Hung (2007), respectively.

A different line of hedging research pursued in the economics literature focuses on constructing a self-financing dynamic portfolio strategy that most closely approximates a given payoff function at a future maturity date. For example, in Heath et al. (2001), Laurent & Pham (1999) and Bertsimas et al. (2001) the authors approach the problem considering minimizing a mean variance or quadratic utility function of the given payoff. They also consider a class of stochastic volatility models in an incomplete financial market. The last two papers use dynamic programming to achieve their goal.

3 The Hedging Setup

We establish the notation and the assumptions we will use for the rest of the paper. The demand for jet fuel occurs at time $T$. It is assumed to be known and is denoted by $d$. The hedging is done using future contracts based either on heating or crude oil and the investor has $N$ trading opportunities prior to $T$. In practice, to avoid actual delivery of the hedging commodity (heating or crude oil), the maturity of the future contract is typically chosen to be one month after $T$. The set of trading dates is given by $T = \{ t_N, \ldots, t_1 \}$. Without loss of generality, we assume that the time between two consecutive trading dates is constant and
is equal to $\Delta$. We also assume that the time between the last trading date $t_1$ and $T$ is also equal to $\Delta$. We assume the risk-free interest rate is known and is equal to $r$.

The price of a future contract at $t$ maturing at $t'$ is denoted by $F_{tt'}$. Moreover, the spot price of oil and the spot price of fuel at $t$ are denoted, respectively, by $P^o_t$ and $P^f_t$. All of them are in dollars per barrel. The vector $W_t = (P^o_t, P^f_t, F_{tt'}, F_{t-\Delta,t'})$, called the information vector, contains the exogenous information relevant to the hedging problem. We discuss in section 6 the models governing the underlying processes.

We denote by $R_t$ the number of future contracts standing prior to the trading decision at time $t$, denoted by $x_t$. Clearly, $R_{t+\Delta} = R_t + x_t$. In the beginning of the hedging period there are no contracts standing, i.e., $R_{t_N} = 0$. Moreover, $x_T = -R_T$, indicating that the future position is closed at the end of the hedging period. For $t \in T$, we impose that $x_t$ must be greater than or equal to $-R_t$, so that the number of standing contracts is never negative. We say that $(W_t, R_t)$ represents the state of our system.

The cash flow at each time period $t \in T \cup \{T\}$ is given by

$$C_{fl}(W_t, R_t, x_t) = 10^3 \left( (F_{tt'} - F_{t-\Delta,t'}) R_t - P^f_T d 1_{\{t=T\}} \right),$$

as the future contract price $F_{tt'}$ is in dollars per barrel and the size of each contract is $10^3$ barrels. Thus, by the end of the hedging period, the realized cost per barrel is given by

$$C^\pi = \sum_{t \in T \cup \{T\}} \frac{e^{r(T-t)} C_{fl}(W_t, R_t, X^\pi_t)}{d},$$

where $X^\pi_t$ is the policy/hedging strategy that determines the trading decision $x_t$ at $t$.

### 4 Hedging Strategies

Before we introduce our policy, we discuss in this section four popular hedging strategies that we use as a comparison in our computational experiments.

- **No-Hedge**: The simplest strategy is to not hedge at all. Using this strategy $x_t = 0$ for $t \in T \cup \{T\}$. Clearly, the realized cost per barrel is $C^{\text{No-hedge}} = -P^f_T$. 
• **Naive**: This strategy is also very straightforward. We buy \( d \) future contracts in the beginning of the hedge period and close the position at \( T \). That is, \( x_{t_N} = d, x_{t_{N-1}} = \cdots = x_t = 0 \) and \( x_T = -d \). In this case, the realized cost per barrel is 
\[
C_{\text{Naive}} = (F_{Tt'} - F_{tN'}d) - P_T^f.
\]

• **Optimal hedge ratio (OHR)**: Also known as the hedge-and-forget ratio, this is the basic strategy that determines the number of contracts to purchase at the beginning of the hedging period based on the ratio given by the covariance between the fuel spot price and the future contract price and the variance of the future contract price. It is optimal when the utility function considered is the variance of the cash flow. We have that 
\[
x_{t_N} = \frac{\text{Cov}(P_{Tt'}, F_{Tt'} - F_{tN't'})}{\text{Var}[F_{Tt'} - F_{tN't'}]} d, \quad x_{t_{N-1}} = \cdots = x_t = 0 \quad \text{and} \quad x_T = -x_{t_N}.
\]

• **Dynamic hedge ratio (OHR\(_t\))**: This is the dynamic version of the previous strategy. At each trading opportunity, new information has become available resulting in an updated covariance and variance. The number of contracts standing is thus rebalanced to reflect the current ratio. In this case, for \( t \in \mathcal{T} \), we have that 
\[
x_t = \frac{\text{Cov}_t(P_{Tt'}, F_{Tt'} - F_{tN't'})}{\text{Var}_t[F_{Tt'} - F_{tN't'}]} d - R_{t-\Delta} \quad \text{and} \quad x_T = -x_t = -R_t.
\]

where the subscript \( t \) indicates that the covariance and the variance are conditional on the information up until time period \( t \).

5 Optimal Dynamic Hedging

We now propose our hedging strategy. We start by defining the utility function under consideration. For \( t \in \mathcal{T} \cup \{T\} \), our utility function is given by
\[
U_t(W_t, R_t, x_t) = C_t^{fl}(W_t, R_t) - u^{vol} \left( C_t^{fl}(W_t, R_t) \right)^2 + 10^3 e^{-r(T-t)} \mathbb{E}_t \left[ (F_{Tt'} - F_{tN'}) (R_t + x_t) - P_T^f d \right] \\
- 10^6 e^{-r(T-t)} u^{vol} \mathbb{E}_t \left[ (F_{Tt'} - F_{tN'}) (R_t + x_t) - P_T^f d \right]^2.
\]
The first term represents the return, while the second term represents its volatility. Moreover, we perform a prospective test on the hedge effectiveness, represented by the last two terms. They measure in terms of return and risk how well the strategy would do by the end of the hedging period, if we were to stop trading at \( t \). Clearly, the test is performed using the price information and the expected values we have available at \( t \). Remember that the rationale behind the \( 10^3 \) and \( 10^6 \) coefficients is that the future contract price \( F_{tt} \) is in dollars per barrel and the size of each contract is \( 10^3 \) barrels. We point out that \( u_{vol} > 0 \) is the investor-defined weight indicating how much risk he is willing to take.

Our objective is to construct a policy \( X_t(W_t, R_t) \), which is a function of the current state \((W_t, R_t)\), that maximizes (I). The trading decision \( x_t \) is determined by this policy.

We solve the problem using dynamic programming. In this framework, the downstream effects of the decision \( x_t \) taken at time \( t \) is determined by the optimal value function at the next time period \( t + \Delta \). The value functions are determined recursively in a backward fashion. At the end of the hedging period \( T \), it is clear that \( V_T(W_T, R_T) = U_T(W_T, R_T, -R_T) \). Recursively, for \( t \in T \), we have that

\[
V_t(W_t, R_t) = \max_{-R_t \leq x_t} E_t \left[ U_t(W_t, R_t, x_t) + e^{-r\Delta} V_{t+\Delta}(W_{t+\Delta}, R_{t+\Delta} + x_t) \right].
\]  

The resulting optimal policy and the corresponding optimal value function are described in the next theorem. We emphasize that the classic curse of dimensionality associated with dynamic programming formulations could have been an issue here, however Theorem 1 shows that the optimal value functions are quadratic in the number of future contracts standing. Computing the optimal policy only depends on our ability to compute the expected values involved.

**Theorem 1.** For \( t \in T \), the optimal policy with respect to the utility function described in (6) is given by

\[
X_t^*(W_t, R_t) = \max \left( -R_t, X_{t, \text{temp}}^*(W_t, R_t) \right),
\]  

(8)
where

\[
X^\star_{t, \text{temp}}(W_t, R_t) = \frac{\mathbf{E}_t \left[ e^{-r(T-t)} (F_{Tv} - F_{tw}) + e^{-\Delta}(F_{t+\Delta t'} - F_{tw}) \right]}{2 \times 10^3 u^{\text{vol}} \mathbf{E}_t \left[ e^{-r(T-t)} (F_{Tv} - F_{tw})^2 + e^{-r\Delta}(F_{t+\Delta t'} - F_{tw})^2 \right]} + \frac{\mathbf{E}_t \left[ e^{-r(T-t)} (F_{Tv} - F_{tw})P^f_T + e^{-r\Delta}(F_{Tv} - F_{tw})P^f T_{t=t_1} \right]}{2 \times 10^3 u^{\text{vol}} \mathbf{E}_t \left[ e^{-r(T-t)} (F_{Tv} - F_{tw})^2 + e^{-r\Delta}(F_{t+\Delta t'} - F_{tw})^2 \right]} d - R_t.
\]

The corresponding optimal value function is a concave quadratic function of \( R_t \) given by

\[
V_t(W_t, R_t) = -10^6 u^{\text{vol}} (F_{tw} - F_{t-\Delta t'})^2 R_t^2 + 10^3 (F_{tw} - F_{t-\Delta t'}) R_t + v_t^0(W_t),
\]

where \( v_t^0(W_t) \) is a term independent of \( R_t \).

Proof. The proof is by backward induction on \( t \). We start with the base case \( t = t_1 \). By definition, for \( t = t_1 \),

\[
V_t(W_t, R_t) = \max_{-R_t \leq x_t} U_t(W_t, R_t, x_t) + e^{-r\Delta} \mathbf{E}_t \left[ U_T(W_T, R_T + x_t, -(R_t + x_t)) \right].
\]

Plugging in (6) and ignoring the terms that are independent of \( x_t \), it turns out that we have to maximize the concave and quadratic function of \( x_t \) given by

\[
f_t(x_t) = -10^6 u^{\text{vol}} \mathbf{E}_t \left[ e^{-r(T-t)} (F_{Tv} - F_{tw})^2 + e^{-r\Delta}(F_{t+\Delta t'} - F_{tw})^2 \right] x_t^2
\]

\[
+ 10^3 \mathbf{E}_t \left[ e^{-r(T-t)} (F_{Tv} - F_{tw}) + e^{-r\Delta}(F_{t+\Delta t'} - F_{tw}) \right] x_t
\]

\[
+ 10^3 \mathbf{E}_t \left[ e^{-r(T-t)} (F_{Tv} - F_{tw})P^f_T d + e^{-r\Delta}(F_{Tv} - F_{tw})P^f_T d \right] x_t
\]

\[
- 2 \times 10^6 u^{\text{vol}} \mathbf{E}_t \left[ e^{-r(T-t)} (F_{Tv} - F_{tw})^2 R + e^{-r\Delta}(F_{t+\Delta t'} - F_{tw})^2 R \right] x_t,
\]

subject to \( x_t \geq -R_t \).

It is easy to see that the solution to the unconstrained maximization problem is the expression for \( X^\star_{t, \text{temp}}(W_t, R_t) \). The optimal policy is thus obtained as we have to enforce the lower bound \(-R_t\). The quadratic value function is obtained by plugging in the expression for the optimal policy in (10), concluding the base case. The induction step assumes that (9) holds for \( t \in T \). The proof that (8) and (9) hold for \( t - \Delta \) follows the same reasoning as the base case. Namely, we use (7) and the induction hypothesis to derive the concave and quadratic function of \( x_{t-\Delta} \) that has to be maximized. The optimal policy and the resulting value function will follow from there. \( \Box \)
We finish this section by noting that the terms on the policy discounted by $e^{-r(T-t)}$ are due to the prospective test, while the terms discounted by $e^{-r\Delta}$ are due to the effect the decision at $t$ will have at the next trading opportunity.

6 The Price Models

We present the models governing the price processes and we derive expressions to the expected values involved in our optimal policy. Our model choice intended to achieve a good tradeoff between accurately describing the stochastic evolution of the processes and providing analytical expressions for the expected values of interest.

Following the model introduced in Schwartz & Smith (2000), the log of the oil spot price is decomposed into two stochastic factors, namely

$$\ln(P_t) = \chi_t + \xi_t,$$

where $\chi_t$ represents a short-term deviation in prices that are not expected to persist. On the other hand, $\xi_t$ represents the long-term equilibrium price level. Changes in $\xi_t$ indicates fundamental changes that are expected to persist.

The short-term deviations are assumed to follow a Ornstein-Uhlenbeck (OU) process

$$d\chi_t = -\kappa \chi_t dt + \sigma_{\chi} dz_{\chi},$$

while the equilibrium level is assumed to follow a Brownian motion process

$$d\xi_t = \mu_{\xi} dt + \sigma_{\xi} dz_{\xi}.$$

In these models, $dz_{\chi}$ and $dz_{\xi}$ are correlated increments of standard Brownian motion processes with $dz_{\chi}dz_{\xi} = \rho_{\chi \xi} dt$. Moreover, $\kappa$ describes the rate at which the short-term deviations are expected to disappear. In addition, we let $\sigma_{\chi}$ and $\sigma_{\xi}$ be the short term and equilibrium volatilities. Finally, $\mu_{\xi}$ is the equilibrium drift rate.

For $t' > t$, it was shown in Schwartz & Smith (2000) that conditioned on the information up to time $t$, $(\chi_{t'}, \xi_{t'})$ are jointly normally distributed with mean vector and covariance
matrix:

\[
\begin{align*}
\mathbb{E}_t[(\chi_t', \xi_t')] &= \left[ e^{-\kappa(t'-t)} \chi_t + \xi_t + \mu_t(t'-t) \right] \tag{11} \\
C_t[(\chi_t', \xi_t')] &= \begin{bmatrix} (1 - e^{-2\kappa(t'-t)})\sigma^2\chi/2\kappa & (1 - e^{-\kappa(t'-t)})\rho_{\chi\xi}\sigma\chi/\kappa \\ (1 - e^{-\kappa(t'-t)})\rho_{\chi\xi}\sigma\chi/\kappa & \sigma^2\xi(t'-t) \end{bmatrix}. \tag{12}
\end{align*}
\]

We thus have that the logarithm of the oil spot price at \(t'\), given \((\chi_t, \xi_t)\), is normally distributed with mean \(\mu_{oil}^{t't'}\) and variance \(\sigma_{oil}^{2,t't'}\), where

\[
\begin{align*}
\mu_{oil}^{t't'} &= e^{-\kappa(t'-t)}\chi_t + \xi_t + \mu_t(t'-t) \tag{13} \\
\sigma_{oil}^{2,t't'} &= (1 - e^{-2\kappa(t'-t)})\sigma^2\chi/2\kappa + \sigma^2\xi(t'-t) + 2(1 - e^{-\kappa(t'-t)})\rho_{\chi\xi}\sigma\chi/\kappa. \tag{14}
\end{align*}
\]

Therefore, given \((\chi_t, \xi_t)\), \(P^o_{t'}\) is log-normally distributed with mean and variance

\[
\begin{align*}
\mathbb{E}_t[P^o_{t'}] &= e^{\mu_{oil}^{t't'} + 0.5\sigma_{oil}^{2,t't'}} \tag{15} \\
\text{Var}_t[P^o_{t'}] &= \left(e^{\sigma_{oil}^{2,t't'}} - 1\right)e^{2\mu_{oil}^{t't'} + \sigma_{oil}^{2,t't'}}. \tag{16}
\end{align*}
\]

We now discuss the relationship between the spot price of oil and the spot price of fuel. We consider the time varying regression (TVR) model presented in [Bos & Gould (2007)]. The results in that paper indicate that the TVR is a good tradeoff between having a simple model and obtaining good results for the correlation between a pair of prices. In the TVR model, the dynamics between the two spot prices are given by the equations

\[
\begin{align*}
d\beta_t &= \sigma_\beta dz_\beta \tag{17} \\
F^f_t &= \alpha + \beta_t P^o_t + \epsilon_t, \tag{18}
\end{align*}
\]

where \(\sigma_\beta\) is the process volatility and \(dz_\beta\) is an increment of a standard Brownian motion process, which is independent of \(dz_\chi\) and \(dz_\xi\). We have that \(\beta_t\) is a martingale representing the correlation between the prices. Moreover, \(\alpha\) is a constant and \(\epsilon_t\) is a noise term that is normally distributed with mean zero and variance \(\sigma^2_\epsilon\). Clearly, we have that

\[
\mathbb{E}_t[F^f_T] = \alpha + \beta_t \mathbb{E}_t[P^o_T], \tag{19}
\]

where \(\mathbb{E}_t[P^o_T]\) is given by (15) replacing \(t'\) with \(T\).
Using the risk-neutral valuation framework (Duffie (1992)), future prices are equal to the expected future spot price under the risk-neutral measure. Therefore, we now focus on the risk-neutral processes governing the oil spot price.

As before, the short-term deviation follows a OU process and the equilibrium level follows a Brownian motion process, given, respectively, by

\[

d\chi_t = (-\kappa \chi_t - \lambda \chi_t)dt + \sigma_\chi dz^\ast_\chi, \\

d\xi_t = (\mu_\xi - \lambda \xi_t)dt + \sigma_\xi dz^\ast_\xi, \tag{20}
\]

where, again, \(dz^\ast_\chi\) and \(dz^\ast_\xi\) are increments of standard Brownian motion processes under the risk-neutral measure with \(dz^\ast_\chi dz^\ast_\xi = \rho \chi \xi dt\). Here, \(\lambda_\chi\) and \(\lambda_\xi\) are, respectively, the short term and the equilibrium risk premiums. Note that the risk-neutral OU process reverts to \(-\lambda_\chi / \kappa\) (instead of 0 in the real world OU process) and the risk-neutral Brownian motion drift is \(\mu^*_\xi \equiv \mu_\xi - \lambda_\xi\) (instead of \(\mu_\xi\)).

Hence, under the risk-neutral measure, conditioned on the information up to time \(t\), \((\chi_t, \xi_t)\) are jointly normally distributed with mean vector and covariance matrix

\[

E^*_t[(\chi_t', \xi_t')] = \left[e^{-\kappa(t'-t)} \chi_t - \left(1 - e^{-\kappa(t'-t)}\right) \frac{\lambda_\chi}{\kappa}, \xi_t + \mu^*_\xi (t' - t)\right] \\

C^*_t[(\chi_t', \xi_t')] = C_t[(\chi_t, \xi_t)].
\]

We use asterisks to denote expectations and (co)variance taken with respect to the risk neutral measure.

We conclude that given \((\chi_t, \xi_t)\), the logarithm of the spot price at \(t'\) under the risk neutral measure is normally distributed with mean and variance

\[

E^*_t[\ln(P^0_{t'})] = e^{-\kappa(t'-t)} \chi_t - \left(1 - e^{-\kappa(t'-t)}\right) \lambda_\chi / \kappa + \xi_t + \mu^*_\xi (t' - t) \\

Var^*_t[\ln(P^0_{t'})] = \sigma^2_{oil_{t'}}. \tag{22}
\]

Since, by definition, \(F_{TT'} = E^*_T[P^0_{t'}]\), we have that

\[

\ln(F_{TT'}) = \ln(E^*_T[P^0_{t'}]) = E^*_T[\ln(P^0_{t'})] + \frac{1}{2} Var^*_T[\ln(P^0_{t'})] \\

= e^{-\kappa(t'-T)} \chi_T + \xi_T + A(t' - T),
\]
where \( A(t) = \mu^*_t t - (1 - e^{-\kappa t}) \frac{\lambda}{\kappa} + \frac{1}{2} \left( (1 - e^{-2\kappa t}) \frac{\sigma^2_\xi}{2\kappa} + \sigma^2_\xi t + 2(1 - e^{-\kappa t}) \frac{\rho_\xi \sigma_\chi \sigma_\xi}{\kappa} \right). \)

We compute some expectations involving \( P_{T'}^t \) and \( F_{T'} \). We start with the conditional expected value and variance of \( \ln(F_{T'}) \) at \( t \), denoted by \( \mu_{t,T'}^{\text{fut}} \) and \( \sigma_{t,T'}^{2,\text{fut}} \), respectively. We have that
\[
\mu_{t,T'}^{\text{fut}} = \mathbf{E}_t \left[ e^{-\kappa(t'-T)} \chi_T + \xi_T + A(t'-T) \right] = e^{-\kappa(t'-t)} \chi_t + \xi_t + \mu_\xi (T-t) + A(t'-T)
\]

and
\[
\sigma_{t,T'}^{2,\text{fut}} = \text{Var}_t \left[ e^{-\kappa(t'-T)} \chi_T + \xi_T + A(t'-T) \right] = e^{-2\kappa(t'-T)} \left( 1 - e^{-2\kappa(T-t)} \right) \frac{\sigma^2_\xi}{2\kappa} + \sigma^2_\xi (T-t) + 2e^{-\kappa(t'-t)} (1 - e^{-2\kappa(T-t)}) \frac{\rho_\xi \sigma_\chi \sigma_\xi}{\kappa}.
\]

Therefore, as with the oil spot price, conditioned on the information until \( t \), \( F_{T'} \) is log-normally distributed with mean and variance
\[
\begin{align*}
\mathbf{E}_t[F_{T'}] &= e^{\mu_{t,T'}^{\text{fut}} + \frac{1}{2} \sigma_{t,T'}^{2,\text{fut}}} \\
\text{Var}_t[F_{T'}] &= \left( e^{\sigma_{t,T'}^{2,\text{fut}}} - 1 \right) e^{2\mu_{t,T'}^{\text{fut}} + \sigma_{t,T'}^{2,\text{fut}}}.
\end{align*}
\]

The conditional expectation of the product between the spot and the future contract price for oil at \( T \) is our next computation. We have that
\[
P_T^o F_{T'} = \exp (\chi_T + \xi_T) \exp \left( e^{-\kappa(t'-T)} \chi_T + \xi_T + A(t'-T) \right) = \exp \left( \left( 1 + e^{-\kappa(t'-T)} \right) \chi_T + 2\xi_T + A(t'-T) \right).
\]

Hence,
\[
\mathbf{E}_t [P_T^o F_{T'}] = \exp \left[ \left( 1 + e^{-\kappa(t'-T)} \right) \mathbf{E}_t[\chi_T] + 2\mathbf{E}_t[\xi_T] + A(t'-T) \right]
+ \frac{1}{2} \left( \left( 1 + e^{-\kappa(t'-T)} \right)^2 \text{Var}_t[\chi_T] + 4\text{Var}_t[\xi_T] \right)
+ 2 \left( 1 + e^{-\kappa(t'-T)} \right) \text{Cov}_t(\chi_T, \xi_T).
\]

We close the section pointing out that
\[
\mathbf{E}_t \left[ P_T^f F_{T'} \right] = \mathbf{E}_t \left[ (\alpha + \beta_T P_T^o + \epsilon_T) F_{T'} \right] = \alpha \mathbf{E}_t [F_{T'}] + \beta_t \mathbf{E}_t [P_T^o F_{T'}],
\]
as \( \beta_T \) is independent of \( P_T^o F_{T'} \), \( \epsilon_T \) is independent of \( F_{T'} \), \( \beta_t \) is a martingale and \( \mathbf{E}_t [\epsilon_T] = 0. \)
7 Processes and Parameter Estimation

The processes \( \chi_t \) and \( \xi_t \) are unobservable, as we only have access to future and spot prices. The same is true for \( \beta_t \), the process governing the correlation between the oil and fuel spot prices. Moreover, the parameters for the oil spot price \( \theta = (\kappa, \sigma_\chi, \sigma_\xi, \lambda_\chi, \mu_\chi^*, \mu_\xi) \) and the parameters \( \alpha, \sigma_\beta^2 \) and \( \sigma_\eta^2 \) for the relationship between oil and fuel prices are also unknown. We thus combine Kalman filter (KF) and maximum likelihood estimation to obtain estimates for these processes and parameters.

The Kalman filter is a recursive procedure to compute estimates of the unobserved processes based on observations, in our case of future and spot prices, that are driven by these processes. It generates updated posterior distributions for the considered processes following the Bayes rule. As is the norm in a Bayesian setting, a prior distribution for the initial value of the unobserved processes is required. One way to view Kalman filtering is to think of it as an updating procedure that consists of forming a preliminary guess about the unobserved processes and then adding a correction to this guess, the correction being determined by how well the guess has performed in predicting the next observation.

We first discuss the estimation of the processes \( \chi_t \) and \( \xi_t \) and the corresponding parameters, represented by \( \theta \). A set of future contracts with different maturity dates is used in the estimation. Using Kalman filter terminology, the transition equation, which is a discrete time version of the stochastic processes governing the oil spot price, models the evolution of the process over time. It is described by

\[
\begin{bmatrix}
\chi_{t+\Delta^v} \\
\xi_{t+\Delta^v}
\end{bmatrix} = g^{tr}\begin{bmatrix}
\chi_t \\
\xi_t
\end{bmatrix} + G^{tr}\begin{bmatrix}
\chi_t \\
\xi_t
\end{bmatrix} + \omega_t,
\]

where \( \omega_t \) is a 2 dimensional column vector of serially uncorrelated, normally distributed disturbances with zero mean and covariance \( \Sigma^\omega \equiv C_t[(\chi_{t+\Delta^v}, \xi_{t+\Delta^v})] \), as given by (12). Moreover, \( \Delta^v \) is the length of the time steps. Note that \( \Delta^v \) might be different from the interval between two trading opportunities denoted by \( \Delta \). We have that

\[
g^{tr} = \begin{bmatrix}
0 \\
\mu_\xi \Delta^v
\end{bmatrix} \quad \text{and} \quad G^{tr} = \begin{bmatrix}
e^{-\kappa \Delta^v} & 0 \\
0 & 1
\end{bmatrix}.
\]

Let \( T' = \{t'_1, \ldots, t'_n\} \) be a set of maturity dates. We consider a vector \( y_t \) containing the
logarithm of the future contracts at \( t \) with maturity date in \( T' \). That is,

\[
y_t = \begin{bmatrix} \ln F_{t,t+t'_1} \\ \vdots \\ \ln F_{t,t+t'_n} \end{bmatrix}.
\]

It holds that changes in the price of long-term future contracts gives information about the equilibrium process, while changes in the difference between near and long term future prices gives information about the short-term deviation process. The relationship between the unobserved processes and \( y_t \) is given by the so-called measurement equation

\[
y_t = g^{me} + G^{me} \begin{bmatrix} \chi_t \\ \xi_t \end{bmatrix} + \nu_t, \tag{26}
\]

where \( \nu_t \) is a \( n \)-dimensional column vector of serially uncorrelated, normally distributed measurement errors with mean zero and covariance \( \Sigma^\nu = \text{diag}(\sigma^2_{\nu_1}, \ldots, \sigma^2_{\nu_n}) \). They can represent errors in the reporting price or error in the model’s fit to observed prices. Finally, we have that

\[
g^{me} = \begin{bmatrix} A(t'_1) \\ \vdots \\ A(t'_n) \end{bmatrix} \quad \text{and} \quad G^{me} = \begin{bmatrix} e^{-\kappa t'_1} & 1 \\ \vdots & \vdots \\ e^{-\kappa t'_n} & 1 \end{bmatrix}.
\]

Assuming that \((\chi_0, \xi_0)\) is normally distributed with mean \( \hat{\mu}_0 = (\hat{\chi}_0, \hat{\xi}_0) \) and covariance matrix \( \hat{\Sigma}_0^{\chi\xi} \), then, using (25)-(26) and a set of observed future prices, the Kalman filter recursively determines that \((\chi_t, \xi_t)\) is normally distributed with mean \( \hat{\mu}_t = (\hat{\chi}_t, \hat{\xi}_t) \) and covariance matrix \( \hat{\Sigma}_t^{\chi\xi} \). It uses the following set of equations:

\[
\begin{align*}
\mu^p_t &= g^{tr} + G^{tr} \hat{\mu}_{t-\Delta^tr} \quad \text{and} \quad \Sigma^p_t = G^{tr} \hat{\Sigma}_0^{\chi\xi} (G^{tr})' + \Sigma^\omega \\
\mu^y_t &= g^{me} + G^{me} \hat{\mu}^p_t \quad \text{and} \quad \Sigma^y_t = G^{me} \Sigma^p_t (G^{me})' + \Sigma^\nu \\
D_t &= \Sigma^p_t (G^{me})'(\Sigma^p_t)^{-1} \\
\hat{\mu}_t &= \mu^p_t + D_t(y_t - \mu^y_t) \quad \text{and} \quad \hat{\Sigma}_t^{\chi\xi} = \Sigma^p_t - D_t \Sigma^p_t D_t'.
\end{align*}
\]

The equations in (27) determine the mean and the covariance matrix of the prior distribution for \((\chi_t, \xi_t)\). That is, given the future price observations until \( t - \Delta^tr \), \((\chi_t, \xi_t)\) is normally distributed with mean \( \mu^p_t \) and covariance matrix \( \Sigma^p_t \). Conditioned on the information up until \( t - \Delta^tr \), we have that \( y_t \) is normally distributed with mean and covariance matrix
given by (28), that is, $\mu_y$ and $\Sigma_y$ determines the likelihood of $y_t$. Finally, (29) represents a correction factor based on the difference between the actual and the predicted value of $y_t$. This factor is used in (30) to update the prior distribution of $(\chi_t, \xi_t)$, obtaining the estimates $\hat{\mu}_t = (\hat{\chi}_t, \hat{\xi}_t)$ and $\hat{\Sigma}_{\chi\xi}^t$.

The Kalman filter equations (25)–(30) assume that the parameters $\theta$ and $(\sigma^2_{\nu_1}, \ldots, \sigma^2_{\nu_n})$ are known, which might not be the case. However, they can be estimated by maximizing the log-likelihood function of $y_{\Delta^r}, \ldots, y_H$, which is given by

$$L_{\text{like}}(y, \mu_y, \Sigma_y) \propto -\frac{1}{2} \sum_{t=\Delta^r}^H \ln \left| \det (\Sigma_y^t) \right| - \frac{1}{2} \sum_{t=\Delta^r}^H (y_t - \mu_y^t)'(\Sigma_y^t)^{-1}(y_t - \mu_y^t), \quad (31)$$

where $H$ is the length of the historical period considered to fit the parameters.

We now focus on the relationship between the oil and the fuel spot prices. It is modeled using the TVR equations (17)-(18). We have that $\beta_t$ is the unobserved process. Using the Kalman filter terminology, the transition equation is given by

$$\beta_{t+\Delta^r} = \beta_t + \eta_t, \quad (32)$$

where $\eta_t$ is a normally distributed disturbance with zero mean and variance $\sigma^2_\eta$. The measurement equation is given by (18). Therefore, assuming a normal prior distribution for $\beta_0$, that is, assuming that $\beta_0$ is normally distributed with mean $\hat{\beta}_0$ and variance $\hat{\sigma}_0^2$, the Kalman filter recursively determines that $\beta_t$ is normally distributed with mean $\hat{\beta}_t$ and variance $\hat{\sigma}_t^2$. For that matter, it uses the observed spot prices for oil and fuel and the following equations:

$$\hat{\beta}^p_t = \hat{\beta}_{t-\Delta^r} \quad \text{and} \quad \hat{\sigma}_t^2 = \hat{\sigma}_{t-\Delta^r}^2 + \sigma^2_\eta, \quad (33)$$

$$\mu^f_t = \alpha + P^o_t + \hat{\beta}^p_t \quad \text{and} \quad \hat{\sigma}_t^2 = (P^o_t)^2 \hat{\sigma}_{t}^2 + \sigma^2_t, \quad (34)$$

$$\hat{\beta}_t = \hat{\beta}^p_t + \left(\frac{\hat{\sigma}_t^2}{\sigma^2_{t} P^o_t}\right) (P^f_t - \mu^f_t) \quad \text{and} \quad \hat{\sigma}_t^2 = \hat{\sigma}_t^2 - \left(\frac{\hat{\sigma}_t^2}{\sigma^2_{t} P^o_t}\right)^2 \hat{\sigma}_t^2, \quad (35)$$

As before, (33) determines the prior distribution of $\beta_t$, while (34) determines the distribution of the fuel spot price given the oil spot price observations until $t - \Delta^r$. Note that $\left(\frac{\hat{\sigma}_t^2}{\sigma^2_{t} P^o_t}\right)$ is the correction factor and (35) determines the distribution of $\beta_t$ given the oil spot price observations until $t$. We use $\hat{\beta}_t$ as an estimator for $\beta_t$. 

16
Again, we might not know the parameters $\alpha$, $\sigma^2_\eta$ and $\sigma^2_\epsilon$. Thus, we maximize the log-likelihood function of $P^f_{\Delta t^v}, \ldots, P^f_{H}$, given by

$$L_{\text{like}}(P^f, \mu^f, \hat{\sigma}^2_{1.f}) \propto -\frac{1}{2} \sum_{t=\Delta t^v}^H \ln(\hat{\sigma}^2_{1.f}) - \sum_{t=\Delta t^v}^H \frac{(P^f_t - \mu^f_t)^2}{2\hat{\sigma}^2_{1.f}}.$$  \hspace{1cm} (36)

8 The Fuel Hedging Algorithm

The computational procedure to solve the jet fuel hedging problem is divided into two main parts. The first part of the HedgeFuel algorithm uses historical data to estimate the unknown parameters and to determine a distribution for the unobservable processes. The second part keeps updating the distribution and takes on the hedging problem itself. It uses the mean of the distribution produced by the Kalman filter as an estimator of the unobservable processes to implement the optimal policy derived in section 5 using dynamic programming. By the end of the algorithm/hedge period, we are able to compute the realized cost per barrel. Figure 1 illustrates the idea.

Figure 1: Combining Kalman filter and dynamic programming.
Figure 2 below describes the first part of the jet fuel hedging algorithm, the historical period. No hedging decision is taken in this part. First, the algorithm requires as input data both the Kalman filter parameters and an initial normal distribution for the processes, see steps 1.0 and 1.1. Then, a set of historical future prices are observed, as in step 1.2. After that, the log-likelihood function (31) is maximized to estimate $\theta$, see step 1.3. Note that as $\sigma_\gamma$, $\sigma_\xi$ and $\Sigma^\nu$ represent standard deviations, we require that they be strictly positive. Moreover, as $\rho_{\gamma\xi}$ represents a correlation, it is restricted to be between $-1$ and $1$. Historical fuel and oil spot prices are observed in steps 1.4, while another maximization procedure takes place in step 1.5 to estimate the parameters related to the correlation between the oil and fuel spot prices. As before, the variables representing standard deviations are constrained to be positive. Finally, the algorithm loops over the historical data and use the Kalman filter equations to return an updated distribution for the processes.

**STEP 1.0:** Set the Kalman filter parameters: $\Delta^{tr}$, $H$ and $T'$.

**STEP 1.1:** Set the prior distribution for $(\chi_0, \xi_0)$ and $\beta_0$.

**STEP 1.2:** Observe $y_{\Delta^{tr}}, \ldots, y_{H\Delta^{tr}}$.

**STEP 1.3:** Determine $\theta = (\kappa, \sigma_\chi, \sigma_\xi, \rho_{\chi\xi}, \lambda_\chi, \mu_\chi^*, \mu_\xi)$ maximizing $L_{\text{like}}(y, \mu^y, \Sigma^y)$ (see (31))

subject to $\sigma_\chi > 0$, $\sigma_\xi > 0$, $-1 \leq \rho_{\gamma\xi} \leq 1$, $\Sigma^\nu \geq 0$.

**STEP 1.4:** Observe $P^o_{\Delta^{tr}}, \ldots, P^o_{H\Delta^{tr}}$ and $P^f_{\Delta^{tr}}, \ldots, P^f_{H\Delta^{tr}}$.

**STEP 1.5:** Determine $(\alpha, \sigma_\eta, \sigma_\xi)$ maximizing $L_{\text{like}}(P^f, \mu^f, \hat{\sigma}^2)$ (see (36))

subject to $\sigma_\eta > 0$, $\sigma_\xi > 0$.

**STEP 1.6:** Do for $t = \Delta^{tr}, 2\Delta^{tr}, \ldots, H\Delta^{tr}$:

**STEP 1.6.a:** Update the distribution of $(\chi_t, \xi_t)$ using (27)-(30)

**STEP 1.6.b:** Update the distribution of $\beta_t$ using (33)-(35)

Figure 2: FuelHedge Algorithm Part 1 - Historical Period

Figure 3 describes the second part of the algorithm, the hedge period. It uses the distribution obtained in the first part as its initial distribution. First, the algorithm requires as
an input the hedge parameters, see step 2.0. It assumes the beginning of the hedge period coincides with the end of the historical period, as \( t_N = H \). Nevertheless, the hedge period can start any time after the historical one. However, the longer the interval, the less accurate the distribution, since it was last updated using prices in the past. Given our assumption that the trade dates are equally spaced, step 2.1 also computes the interval \( \Delta \) between them. We have also assumed that the maturity date of the future contracts is one month after the hedging period and that the initial future position is equal to zero. Step 2.2 initializes the cost per barrel following our policy.

At each trading date, the algorithm observes the current price of the future contract maturing at \( t' \), see step 2.3.a. Using this information and the current mean of the unobservable processes as an estimator, it computes both the expected values required to determine the trading decision and the trading decision itself, see steps 2.3.b and 2.3.c. The decision generates a cashflow, given by (4), that is used to update the cost per barrel, see step 2.3.d. The decision also affects the number of future contracts standing, see step 2.3.e.

In between two trading dates, the algorithm keeps observing future/spots prices and using the Kalman filter equations to update the distribution of the unobservable processes, see steps 2.3.f.i-2.3.f.iii.

Finally, at the end of the hedge period, the algorithm observes the current future and fuel spot prices, see step 2.4. Then, the future position is closed and jet fuel is bought to meet the demand, see step 2.5. The algorithm terminates returning the realized cost per barrel.

9 Numerical Experiments

We compare our strategy to the hedging strategies described in section 4, since these are the ones widely used in practice. We use actual price data to run the experiments.

We start by describing the data. We obtained weekly oil and fuel spot prices from January
STEP 2.0: Set hedge parameters: $T$, $N$, $d$, $p^m$, $p^r$, $u^{vol}$.

STEP 2.1: Set $t_N = H$, $\Delta = T - t_N/N$, $t' = T + 1/12$, $R_{t_N} = 0$.

STEP 2.2: $C^* = 0$.

STEP 2.3: Do for $t = t_N, t_N + \Delta, \ldots, T - \Delta$:

- **STEP 2.3.a:** Observe $F_{t'}$.
- **STEP 2.3.b:** Use $(\hat{\chi}_t, \hat{\xi}_t)$ and $\hat{\beta}_t$ as an estimator for $(\chi_t, \xi_t)$ and $\beta_t$ to compute
  \[ E_t\left[P^f_T\right], \ E_t\left[P^f_T F_{T'}\right], \ E_t\left[F_{T'}\right], \ E_t\left[F_{t+\Delta t'}\right], \ E_t\left[F_{t+2\Delta t'}\right]. \]

- **STEP 2.3.c:** Compute $X_t^\ast(W_t, R_t)$ using (8). Set $x_t = X_t^\ast(W_t, R_t)$.
- **STEP 2.3.d:** $C^* = C^* + e^{r(T-t)}C^f_t(W_t, R_t, x_t)$.
- **STEP 2.3.e:** $R_{t+\Delta} = R_t + x_t$.
- **STEP 2.3.f:** Do for $\tilde{t} = t + \Delta_t, t + 2\Delta_t, \ldots, t + \Delta$:
  - **STEP 2.3.f.ii:** Observe $y_{\tilde{t}}$, $P^f_{\tilde{t}}$, $P^p_{\tilde{t}}$.
  - **STEP 2.3.f.iii:** Update the distribution of $(\chi_{\tilde{t}}, \xi_{\tilde{t}})$ using (27)-(30).
  - **STEP 2.3.f.iii:** Update the distribution of $\beta_{\tilde{t}}$ using (33)-(35).

STEP 2.4: Observe $P^f_T$ and $F_{T'}$.

STEP 2.5: Close future position ($x_T = -R_T$) and buy $10^3d$ barrels of jet fuel.

STEP 2.6: Return cost per barrel $C^* = (C^* + C^f_T(W_T, R_T, x_T))/d$.

---

**Figure 3:** FuelHedge Algorithm Part 2 - Hedge Period

2000 to October 2007 from the Energy Information Administration (EIA) website\(^2\) which is a statistical agency of the U.S. Department of Energy. The price of oil refers to crude oil traded in the spot market at Cushing, Oklahoma. The price of fuel refers to kerosene-type jet fuel traded in the spot market in the New York Harbor. For the same period, we also obtained futures contracts prices from the Datastream database. The future prices refer to NYMEX contracts on light sweet crude oil.

Figure\(^3\) shows some of the specifications for NYMEX future contracts on crude oil. It also shows two graphs: one conveys future contract prices that matures in the front month and

\(^2\)http://www.eia.doe.gov/
The other conveys the spot prices for oil and fuel. In fact, this figure is the best justification for hedging: futures, oil and fuel prices are skyrocketing. Even the margin requirements have been changed for November/2007 to accommodate the high prices. We can also see from this figure that the spot prices for oil and fuel are indeed closely correlated, with a shift separating the two of them.

We now focus on the first part of the HedgeFuel algorithm, where we use historical data to estimate the process parameters and to get the Kalman filter started. Our historical period $H$ goes from January 2000 to October 2006. Moreover, we consider a weekly frequency, that is, $\Delta_t$ is equal to one week. Finally, our vector of future prices $y_t$ contains five contracts, each maturing 1, 4, 8, 12 and 15 months, respectively, from $t$. Thus, $T' = \{1, 4, 8, 12, 15\}$ months. To get the algorithm started, we also have to provide a prior distribution for the processes. We set $(\hat{x}_0, \hat{\xi}_0) = (0, 2)$ and the covariance matrix $\Sigma_{x^\xi}^0$ equal to the identity matrix. Moreover, we set $\hat{\beta}_0 = 1$ and $\hat{\sigma}_{0,0}^2 = .5$. 

Figure 4: NYMEX oil contract description and price data.
Table 1 shows the estimated values for the parameters. The log likelihood functions were maximized using the Matlab function \textit{fminunc}. Figure 5 shows both the actual oil spot prices and the estimated one, using the Kalman filter. We can infer that the two-factor price model did reflect the actual price behavior and the Kalman filter did do a good job tracking the two unobservable processes.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\kappa$</td>
<td>1.7501</td>
</tr>
<tr>
<td>$\sigma_\xi$</td>
<td>0.1813</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>8.1025</td>
</tr>
<tr>
<td>$\sigma_\beta$</td>
<td>0.0001</td>
</tr>
<tr>
<td>$\rho_{\chi \xi}$</td>
<td>0.4032</td>
</tr>
<tr>
<td>$\lambda_\chi$</td>
<td>0.2439</td>
</tr>
<tr>
<td>$\sigma_{\nu_1}$</td>
<td>0.0062</td>
</tr>
<tr>
<td>$\mu_\xi$</td>
<td>0.0120</td>
</tr>
<tr>
<td>$\sigma_{\nu_2}$</td>
<td>0.0100</td>
</tr>
<tr>
<td>$\mu_\xi$</td>
<td>0.0744</td>
</tr>
<tr>
<td>$\sigma_{\nu_3}$</td>
<td>0.0000</td>
</tr>
<tr>
<td>$\sigma_\chi$</td>
<td>0.6429</td>
</tr>
<tr>
<td>$\sigma_{\nu_4}$</td>
<td>0.0125</td>
</tr>
</tbody>
</table>

Table 1: Estimated parameters using actual prices from Jan/2000 to Oct/2006

We close this part by reporting the updated processes distributions. We obtained

$$(\hat{\chi}_H, \hat{\xi}_H) = \begin{bmatrix} -0.1342 \\ 4.2542 \end{bmatrix} \quad \text{and} \quad \hat{\Sigma}_{HH}^{\chi, \xi} = 10^{-3} \begin{bmatrix} 0.2025 & -0.0351 \\ -0.0351 & 0.0061 \end{bmatrix}.$$
Moreover, $\hat{\beta}_H = 1.1286$ and $\hat{\sigma}^2_{H,\beta} = 5.5039^{-11}$. We note that the standard deviations are pretty small, indicating that the estimators (the mean distribution) should be quite reliable.

In all the experiments, we consider the hedge period to be one year, i.e., $T = 1$ year. We also consider four trading opportunities, implying that the time between two consecutive decision dates is three months. Thus, $N = 4$ and $\Delta = 3$ months. We have set the interest rate to zero and the demand to be 100 contracts. We note that picking the day/time to execute the trading order is out of the scope of this work.

First, we want to see how our strategy compares to the other strategies. We consider the price data from January/2000 to the end of October/2006 as our historical period. The hedge period goes from November/2006 to November/2007. It is a simulation exercise, as we want to consider several runs in order to observe the sample volatility of the price per barrel under the different approaches. Thus, we have used the price models introduced in section 6 (considering the parameter values described in table 1) to simulate 300 different sets of weekly prices over the hedging period. Of course, our algorithm still assumed the processes were unobservable and the Kalman filter equations were used to produce the processes estimators. We used the same estimators for the competing policies.

Figure 6 shows the price per barrel produced by the different hedging strategies, when the volatility weight for our approach ranged from .2 to .5. It also shows the 95% upper confidence bound for the prices. Comparing the price produced by the No-Hedge strategy with the actual fuel price for November/2007 conveyed in figure 4, we can see that the simulation produced a realistic outcome. We can also observe that the higher the volatility weight, the smaller the sample volatility. Clearly, the reduction in volatility resulted in a higher mean cost per barrel. Small weight values indicates the investor is emphasizing speculation rather than hedging.

Note that the graph is divided into seven regions. Region I corresponds to very small volatility weights. We can observe the gambling nature of our approach: very low mean price per barrel but substantial risk. In region II, even though the weights are still small and the risks are still large, our approach does better than not hedging at all. In region III, we dominate the hedge-and-forget OHR strategy, providing smaller costs and smaller risks.
On the other hand, there is no clear dominance between our approach and the dynamic OHR approach over regions IV and VI. The investor has to decide what suits his needs best: either a better price per barrel or a smaller risk. Region V shows a narrow range for which the dynamic OHR dominates our approach by small margins. Finally, in region VII we outperform the naive approach.

We can infer from this experiment that considering a more elaborate utility function does pay off. More importantly, the ability to turn a knob on the volatility weight allows the investor to decide which level of price/risk he is more comfortable with.

Figure 7 illustrates the speculative nature of our policy for different weights (using the same sample path). Each figure shows the decision (buying or selling), the number of contracts outstanding, and the final demand, over the horizon.

The strategy follows the same pattern for all the different weights; the difference is the
size of the future position. Since in the current market situation the oil/fuel prices are trending upward, the strategy tells the investor to buy several contracts, more than the actual demand, in the beginning of the hedge period. Then, as the prices reveal themselves along time, the strategy is to rebalance the position selling some of the contracts. Note that the higher the volatility weight, the closer to the actual demand is the number of contracts standing by the end of the period. The high speculative nature of the strategy can be observed when $\nu^{vol} = .2$, as the number of contracts standing is way beyond the demand.

We conclude this section showing the actual price per barrel if we were to follow the different approaches on a weekly basis. As before, the hedging period is one year and there are four trading opportunities. Figure 8 conveys the actual prices for volatility weights equal to .2, .3, .4 and .5. The reported price for a given month/year assumes that the hedging
period was the previous twelve months. Moreover, only prices up to the same month in
the previous year were used as historical data to initialize the Kalman filter. The prices
during the hedging period, were considered progressively by the algorithm, in an on-line
environment. Note that we are not able to provide an upper bound limit for the price, since
we only had one sample path, the actual prices.

![Figure 8: Actual price per barrel](image)

We can infer from figure 8 that our approach outperformed, for most weeks, the other
approaches for the two smallest weights. In particular, for $u^{vol} = .2$, our approach generated
negative cost for a couple of weeks, indicating that the speculation inherent in the hedging
strategy actually made money. In general, the actual cost per barrel closely followed the
results obtained when the prices for the hedge period were simulated.
10 Conclusions and Further Research

We proposed a novel trading strategy for the jet fuel hedging problem using oil future contracts. Our policy is optimal with respect to a utility function that considers risk, return and a prospective test on the hedge effectiveness. Moreover, different levels of risk aversion are taken into consideration through a parameter set by the investor. On the contrary, the most traditional strategy is the one that only minimizes the risk.

We gathered actual price data and compared our strategy to other well established approaches. We were able to observe that it dominates the others for certain risk levels. In addition to that, the algorithmic procedure to compute our policy, combining dynamic programming and Kalman filter, is very practical and easy to implement. We can thus conclude that our approach can be quite useful for practitioners.

We also analyzed the nature of our policy. As expected, if the tolerance for risk is high, it has a speculative nature. On the other hand, the more risk averse, the more the policy concentrates only on hedging, generating higher but more stable costs per barrel.

A future direction for research is the consideration of transaction costs in the utility function, as trading future contracts usually involves brokerage commissions and margin requirements. However, when these costs are incorporated in our utility function, the corresponding dynamic program can not be solved analytically, leading to the curse of dimensionality. It would be necessary to consider a technique such as approximate dynamic programming to produce a trading policy that reflects such costs.

References


