

# Mean-variance hedging with oil futures

Liao Wang      Johannes Wissel

School of ORIE, Cornell University  
Ithaca, NY 14853  
USA  
{lw432, jw674}@cornell.edu

27th August 2011

## Abstract

We analyze mean-variance-optimal dynamic hedging strategies in oil futures for oil producers and consumers. In a model for the oil spot and futures market with Gaussian convenience yield curves and a stochastic market price of risk, we find analytical solutions for the optimal trading strategies. An implementation of our strategies in an out-of-sample test on market data shows that the hedging strategies improve long-term return-risk profiles of both the producer and the consumer.

**Key words** mean-variance hedging, fuel hedging, energy futures market

**MSC 2010 Classification Numbers** 60H30, 91B30, 91G20, 91G80

**JEL Classification Numbers** C61, G11, G13

## 1 Introduction

Hedging of financial market risk is an important problem in operations research. Consider a non-financial corporation which is exposed to commodity price risk. Suppose that our corporation consumes at a constant rate of one unit of commodity per time, that it is not able to store the commodity in significant amounts, and thus has to continuously purchase the commodity on the spot market for immediate consumption. The cumulative discounted cashflow arising from consumption during the planning period  $[0, T]$  is then given by

$$H_T = - \int_0^T e^{-ru} S_u du,$$

where  $r$  denotes the (constant) interest rate and  $S_t$  denotes the commodity spot price at time  $t$ . If there exists a futures contract for the commodity with futures price  $F(t, u)$  at time  $t$  and delivery date  $u$ , then  $F(u, u) = S_u$  and the random cashflow  $H_T$  can be perfectly hedged at time 0 by entering into  $du$  futures contracts  $F(0, u)$  for each  $u \in [0, T]$ . In practice, this is usually an unrealistic assumption. First, there might not exist a futures contract for each delivery date  $u$ . Secondly, and more importantly, there are commodities for which there does not exist a liquid futures market. In this case, it is not possible to replicate the cashflow  $H_T$  by trading in the financial market. Nevertheless, it is often possible to *partially* hedge the price risk by trading in futures contracts on a *substitute* commodity whose returns are correlated with the changes in the spot price  $S_t$ .

A typical example is the case of jet fuel (kerosene), for which there does not exist a futures market. To manage their exposure to jet fuel price risk, some airlines use futures on crude oil or heating oil as a proxy to hedge the price risk. If liquidly traded futures contracts  $F(t, u)$  are available for maturity dates  $u = T_1, \dots, T_m$ , the airline can employ a self-financing trading strategy in the futures market by taking positions  $\vartheta_t^j$  in the contracts  $F(t, T_j)$  at time  $t$ , which generates a cumulative discounted gains process

$$G_T(\vartheta) = \sum_{j=1}^m \int_0^T \vartheta_u^j e^{-ru} dF(u, T_j)$$

at time  $T$ . If the initial wealth is  $X_0$ , the discounted terminal wealth with hedging strategy  $\vartheta_t^j$  is given by

$$X_T(\vartheta) = X_0 + G_T(\vartheta) + H_T.$$

In general, it is impossible to find a deterministic initial wealth  $X_0$  and a predictable trading strategy  $\vartheta$  such that  $X_T(\vartheta)$  is zero, that is, the kerosene price risk during the planning horizon  $[0, T]$  cannot be perfectly hedged at time 0 by trading in the financial market. In other words, we are faced with a hedging problem in an *incomplete market*.

There are both theoretical arguments and empirical evidence that financial hedging can increase firm values; we refer to Bertus et al. [1] for an overview on literature on these questions. In this paper, we are mainly interested in analyzing optimization problems for the hedging strategy. Much of the applied work on optimization problems in incomplete markets uses *quadratic criteria* because of their analytical tractability. Moreover, most of the literature on fuel hedging only deals with the problem of *minimizing the variance* of the hedging error as in [1], and concentrates on one-period hedging strategies, see for instance [12] and [4] for overviews on various futures hedging approaches. One exception is the paper by Nascimento and Powell [14], who introduce a *mean-variance tradeoff* into their optimality criterion, and consider dynamic trading strategies. Our approach is closest in spirit to [14]. We consider the three Markowitz-type problems

$$U(a) = \sup_{\vartheta \in \mathcal{A}} \left( E[X_T(\vartheta)] - a \text{Var}[X_T(\vartheta)] \right), \quad (1.1)$$

$$B(m) = \inf_{\vartheta \in \mathcal{A}} \left\{ \text{Var}[X_T(\vartheta)] \mid E[X_T(\vartheta)] = m \right\}, \quad (1.2)$$

$$C(v) = \sup_{\vartheta \in \mathcal{A}} \left\{ E[X_T(\vartheta)] \mid \text{Var}[X_T(\vartheta)] = v \right\} \quad (1.3)$$

over a suitable set  $\mathcal{A}$  of admissible trading strategies for given risk aversion coefficient  $a > 0$ , target return  $m \in \mathbb{R}$ , and target variance  $v > 0$ . This approach allows us to compare, quantitatively and via analytical formulas, the performance of an optimal hedging strategy with the case of an agent who does not employ financial hedging strategies, via economically intuitive choices for the coefficients  $m$  and  $v$ . Indeed, the hedger may choose either the target return  $m$  or the target variance  $v$  to be equal to the non-hedger's profile, and solve the corresponding problem (1.2) for the optimal variance or (1.3) for the optimal expectation from hedging. Nascimento and Powell [14] consider a criterion similar to (but more complex than) (1.1) using an additive quadratic utility function which is tailor-made to allow for a tractable solution via a dynamic program. In their setup, the quantitative interpretation (and hence finding a reasonable numerical value) of the risk aversion coefficient seems to be less obvious.

We use an exponentially-affine version of the Miltersen and Schwartz [13] market model for the term structure of futures prices to model the WTI crude oil futures market, which appears (in various equivalent formulations) in many articles such as [3], [14], [1]. In this model, the futures prices  $F(t, T)$  are of the

form

$$F(t, T) = F_t e^{(T-t)r - a(T-t) - b(T-t)X_t}, \quad T \geq t,$$

where  $a(\cdot), b(\cdot)$  are deterministic functions with  $a(0) = b(0) = 0$ , and  $(\log F_t, X_t)$  is a two-dimensional Gaussian Itô process whose value at time  $t$  can be observed from the futures term structure  $F(t, T)$ . We can interpret  $F_t$  as the futures price for immediate delivery, or equivalently the crude oil spot price, and  $\frac{a(T-t) + b(T-t)X_t}{T-t}$  as the convenience yield of physical crude oil for the time period  $[t, T]$ . Under a suitable choice of the functions  $a(\cdot), b(\cdot)$  and a market price of risk which is affine in  $X_t$ , the futures market model can be shown to be arbitrage-free. Finally, we assume that the kerosene spot price  $S_t$  contains a risk component which is independent of the filtration generated by the crude oil futures price processes. Therefore the price  $S_t$  (and thus the cashflow  $H_T$ ) cannot be replicated by trading in the financial instruments, i.e., the futures contracts  $F(t, T)$ , and the combined futures and spot market model is incomplete.

The Markowitz problems (1.1) – (1.3) are closely related to the *quadratic hedging problem* of minimizing the expected quadratic hedging error (see (3.6) below), which has been extensively studied in various setups and levels of generality by many articles such as [18], [11], [16] and [5], to name but a few. Although the *general structure* of the solution to this problem is well-understood, obtaining closed form expressions for the *solution in a concrete model* can be quite tricky in applications. In particular, this task usually becomes difficult if the market price of risk is not *deterministic*, or at least *bounded*. While these assumptions are not satisfied in our model, in which the market price of risk follows an Ornstein-Uhlenbeck process, we are able to solve the quadratic hedging problem and consequently the optimization problems (1.1) – (1.3) in closed form, by exploiting the exponentially-affine Gaussian structure of the price processes.

The paper is organized as follows. We start by describing the WTI futures and kerosene spot market and introducing the market model in section 2. Our main results are presented in section 3, where we formulate the optimization problems, and apply general results from the theory of quadratic hedging to solve these problems within our model in closed form. Section 4 describes the calibration and numerical tests of our model using market data. We summarize our conclusions in section 5. Finally, the proofs of the main results are provided in the appendix.

## 2 Model setup

In this section, we describe the exponentially-affine Gaussian model for the futures market within the context of the Miltersen and Schwartz [13] market model for the term structure of futures prices and convenience yields.

### 2.1 The market

We consider a market consisting of a riskless asset with price  $B_t$ , a family of futures contracts with maturity dates  $T$  and futures prices  $F(t, T)$ , and a commodity with spot price  $S_t$  at time  $t$ . In our applications,  $F(t, T)$  will be futures prices on WTI crude oil, and  $S_t$  will be either

- the spot price of the underlying commodity (WTI for immediate delivery at Cushing, OK), or
- the spot price of a derivate of the underlying commodity (typically an oil refinery product for which futures contracts do not exist, e.g. jet fuel).

In the WTI futures market, for every month until a certain time horizon there exists a NYMEX traded futures contract for delivery in that month, which has a maturity (i.e., end of trading) date given by a specific day in the prior month (usually three business days before the 25th calendar day, see the EIA homepage for detailed definitions). At each time  $t$ , the futures contracts with the next  $m$  maturity dates are called contracts 1 through  $m$ , and we denote their maturity dates by  $T_j = T_j(t)$  ( $j = 1, \dots, m$ ). We use the notation  $T_j$  for the maturity date of contract  $j$ , keeping in mind that  $T_j = T_j(t)$  by definition depends on  $t$ . The maturity date  $T_1$  of contract 1 at time  $t$  is the earliest maturity date for which a contract is available at time  $t$ . When contract 1 expires, contracts 2 through  $m$  become the new contracts 1 to  $m - 1$ .

## 2.2 Stochastic models for spot price and futures term structures

For arbitrage analysis, it is convenient to assume that at each time  $t$ , there are futures contracts traded for all maturities  $T \in [t, T^*]$  for some fixed time horizon  $T^* > 0$ . We assume that we can continuously trade in the futures contracts  $F(t, T)$  and the riskless asset  $B_t$  without transaction costs, but trading in the physical commodity is limited to buying or selling for instantaneous consumption or delivery.

We assume that the interest rate on the riskless asset is constant; this can easily be generalized to deterministic time-dependent interest rates. The rationale for this assumption is that volatility in oil spot prices and futures term structures is significantly higher than interest rate volatility, and hence our focus is on price risk rather than interest rate risk. Hence we have  $B_t = e^{rt}$ . The futures and spot market are modeled as follows. We define the spread between the logarithms of the commodity spot price and the discounted contract 1 futures price by

$$Y_t = \log S_t - \log \left( F(t, T_1) e^{-(T_1-t)r} \right). \quad (2.1)$$

We express the futures prices as

$$F(t, T) = F_t e^{(T-t)r - \int_t^T \epsilon(t, s) ds}, \quad T \in [t, T^*], \quad (2.2)$$

where  $F_t$  can be interpreted as the price of a hypothetical futures contract on WTI for immediate delivery. If  $S_t$  is the WTI spot price and  $F_t = S_t$ , we can interpret  $\frac{1}{T-t} \int_t^T \epsilon(t, s) ds$  as the convenience yield of physical crude oil for the time period  $[t, T]$ . However, we do not require that  $F_t = S_t$ , and this relation clearly will not hold if  $S_t$  is not the underlying WTI spot price, but the spot price of some crude oil derivate. Indeed, both  $F_t$  and  $\epsilon(t, T)$  are not direct market observables, but quantities that are deduced from the futures market price data  $F(t, T_j)$  and the interest rate  $r$  (see section 4.1). In Figure 1, we plot the historical values of the spread process  $Y_t$  if  $S_t$  is the spot price (in \$ per gallon) of jet fuel and  $F(t, T_1)$  is the futures price (in \$ per barrel) of WTI.

The dynamics of the spot and futures prices are now modeled as follows. Let  $(\Omega, \mathcal{F}, \mathbb{F}, P)$  be a filtered probability space carrying a Brownian motion  $W_t = (W_t^0, \dots, W_t^{d+1})$  with  $d+2$  independent components, and  $\mathcal{F} = \mathcal{F}_{T^*}$  with  $\mathbb{F} = (\mathcal{F}_t)_{t \in [0, T^*]}$  the  $P$ -augmented filtration generated by the Brownian motion  $W_t$  on  $[0, T^*]$ . Let  $W_t^E = (W_t^1, \dots, W_t^d)$  and  $W_t^F = \sum_{i=0}^d \rho_i W_t^i$  for correlation coefficients satisfying  $\sum_{i=0}^d \rho_i^2 = 1$ . We assume that  $F_t$  and  $\epsilon(t, T)$  are adapted stochastic processes with dynamics

$$dF_t = F_t \mu_t dt + F_t \sigma_t dW_t^F, \quad (2.3)$$

$$d\epsilon(t, T) = \alpha(t, T) dt + \sigma_\epsilon(t, T) \cdot dW_t^E, \quad (2.4)$$

where  $\mu_t$ ,  $\sigma_t$  and  $\alpha(t, T)$  are 1-dimensional and  $\sigma_\epsilon(t, T)$  is a  $d$ -dimensional predictable process satisfying the usual integrability conditions. Finally the spread  $Y_t$  is given by a continuous  $\mathbb{F}$ -adapted process.

### 2.3 Absence of arbitrage

Throughout this paper, we assume that there is no arbitrage in the riskless asset and futures market. By the fundamental theorem of asset pricing, this is essentially equivalent to the existence of an equivalent local martingale measure for the futures prices. The following result characterizes arbitrage-free futures term structure models in a similar spirit as the HJM framework for interest rate term structures.

**Theorem 2.1.** *There exists a local martingale measure  $Q \approx P$  on  $\mathcal{F}$  for all futures price processes  $F(t, T)$ ,  $T \leq T^*$  if and only if there exists a  $(d+2)$ -dimensional market price of risk process  $\theta_t = (\theta_t^0, \theta_t^E, \theta_t^{d+1})$  with  $\theta^E = (\theta^1, \dots, \theta^d)$  such that*

$$\alpha(t, T) = \sigma_\epsilon(t, T) \cdot \left( \theta_t^E + \int_t^T \sigma_\epsilon(t, s) ds - \sigma_t \rho \right), \quad (2.5)$$

$$\theta_t^0 \rho_0 + \theta_t^E \cdot \rho = \frac{\mu_t - r + \epsilon(t, t)}{\sigma_t}, \quad (2.6)$$

where  $\rho = (\rho_1, \dots, \rho_d)$ . In this case  $Q$  is of the form

$$\left. \frac{dQ}{dP} \right|_{\mathcal{F}_t} = e^{-\int_0^t \theta_u \cdot dW_u - \frac{1}{2} \int_0^t \|\theta_u\|^2 du}, \quad t \in [0, T^*].$$

In particular, the process  $W_t + \int_0^t \theta_u du$  is a  $(d+2)$ -dimensional  $Q$ -Brownian motion.

*Proof.* This follows by taking deterministic interest rates in the model of section III in Miltersen and Schwartz [13], and changing from the risk-neutral to the real world measure.  $\blacksquare$

**Remark.** For any pair  $(\theta_t^0, \theta_t^E)$  satisfying (2.5), (2.6), there exist infinitely many equivalent local martingale measures, parametrized by the market price of risk component  $\theta_t^{d+1}$ , so the market is incomplete. In typical applications,  $Y_t$  and thus  $S_t$  will depend on  $W_t^{d+1}$ . Since  $S_t$  is not a traded asset, an  $\mathcal{F}$ -measurable random variable in general cannot be replicated by trading in the futures contracts  $F(t, T)$ .

### 2.4 Affine convenience yield curve models

To obtain an analytically tractable model, we will from now focus on a Gaussian framework as considered various papers such as [3], [14], [1], and [13]. Let  $d = 1$ , so  $W_t = (W_t^0, W_t^1, W_t^2)$  is a 3-dimensional Brownian motion on  $(\Omega, \mathcal{F}, \mathbb{F}, P)$ , with  $\mathcal{F} = \mathcal{F}_{T^*}$  and  $\mathbb{F}$  generated by  $(W_t)_{t \in [0, T^*]}$ . We assume that  $\sigma_t = \sigma$  in (2.3) is constant, and the convenience yield curve is of the affine form

$$\int_t^T \epsilon(t, s) ds = a(T-t) + b(T-t)X_t \quad (2.7)$$

with  $X_t$  following a centralized Ornstein-Uhlenbeck process

$$dX_t = -\kappa X_t dt + \eta dW_t^1 \quad (2.8)$$

for constants  $\kappa, \eta > 0$  and  $a(\cdot)$  and  $b(\cdot)$  deterministic functions satisfying  $a(0) = b(0) = 0$  and  $b'(0) = 1$ . Futures prices are then given by

$$F(t, T) = F_t e^{(T-t)r - a(T-t) - b(T-t)X_t}, \quad (2.9)$$

$$dF_t = F_t \mu_t dt + F_t \sigma (\rho_0 dW_t^0 + \rho_1 dW_t^1) \quad (2.10)$$

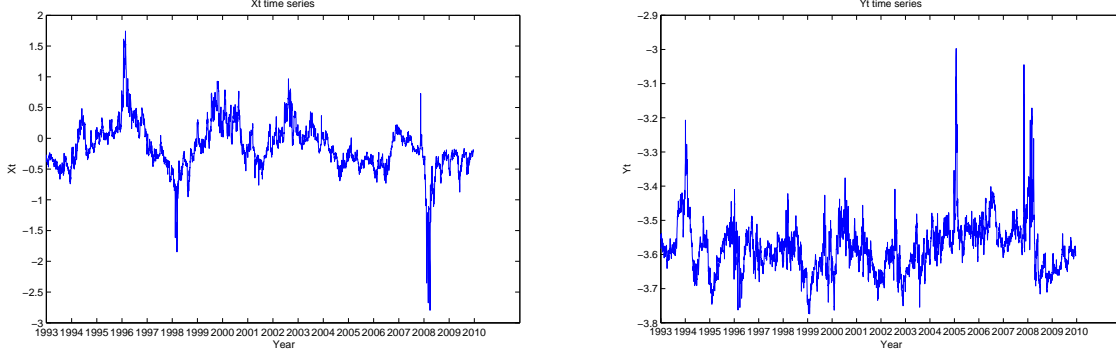


Figure 1: Historical values of the processes  $X_t$  and  $Y_t$ . Both time series exhibit mean-reversion behavior.

with  $\rho_0 = \sqrt{1 - \rho^2}$  and  $\rho_1 = \rho$  for some  $\rho \in (-1, 1)$ . Finally we suppose that the components  $(\theta^0, \theta^1)$  of the market price of risk in Theorem 2.1 have the form

$$\hat{\theta}_t^i = \beta_i + \gamma_i X_t, \quad i = 0, 1 \quad (2.11)$$

for constants  $\beta_i$  and  $\gamma_i$ ,  $i = 0, 1$ . In Figure 1, we plot the historical values of  $X_t$  implied by the futures contracts 1 to 4 (see section 4.1 for details how  $X_t$  is implied from the futures term structure).

**Remark.** The structure of the model (2.7), (2.8) is preserved under an affine transformation of the process  $X_t$ . The conditions that  $X_t$  is centralized and  $b'(0) = 1$  are therefore imposed without loss of generality to ensure a canonical form with a minimal number of parameters.  $\diamond$

**Theorem 2.2.** *Suppose that we have an arbitrage-free futures market model as in Theorem 2.1. Assumptions (2.8) – (2.11) then imply that*

$$a(\tau) = \epsilon_0 \tau + \epsilon_1 e_2(\beta, \tau) - \frac{\eta^2}{2} e_3(\beta, \tau), \quad (2.12)$$

$$b(\tau) = e_1(\beta, \tau) \quad (2.13)$$

with  $\beta = \kappa + \gamma_1 \eta$  and constants  $\epsilon_0, \epsilon_1$ , where the functions  $e_1, e_2, e_3$  are defined by

$$\begin{aligned} e_1(\beta, \tau) &= \int_0^\tau e^{-\beta s} ds &&= \begin{cases} \frac{1}{\beta}(1 - e^{-\beta\tau}) & (\beta \neq 0) \\ \tau & (\beta = 0) \end{cases} \\ e_2(\beta, \tau) &= \int_0^\tau e_1(\beta, s) ds &&= \begin{cases} \frac{1}{\beta} \left( \tau - \frac{1}{\beta}(1 - e^{-\beta\tau}) \right) & (\beta \neq 0) \\ \frac{\tau^2}{2} & (\beta = 0) \end{cases} \\ e_3(\beta, \tau) &= \int_0^\tau (e_1(\beta, s))^2 ds &&= \begin{cases} \frac{1}{\beta^2} \left( \tau - \frac{2}{\beta}(1 - e^{-\beta\tau}) + \frac{1}{2\beta}(1 - e^{-2\beta\tau}) \right) & (\beta \neq 0) \\ \frac{\tau^3}{3} & (\beta = 0). \end{cases} \end{aligned}$$

Moreover

$$\epsilon(t, t) = \epsilon_0 + X_t,$$

and the futures price dynamics are given by

$$dF(t, T) = F(t, T) \left( \sigma \rho_0 (dW_t^0 + \hat{\theta}_t^0 dt) + (\sigma \rho_1 - \eta e_1(\beta, T - t)) (dW_t^1 + \hat{\theta}_t^1 dt) \right). \quad (2.14)$$

*Proof.* This follows from the results on affine term structure models in [9]. For completeness, we give the proof in the appendix.  $\blacksquare$

Finally, we shall also model  $Y_t$  as an Ornstein-Uhlenbeck process satisfying

$$dY_t = \phi(b - Y_t)dt + \nu(c_0dW_t^0 + c_1dW_t^1 + c_2dW_t^2), \quad (2.15)$$

with  $c_0, c_1, c_2 \in [-1, 1]$  and  $c_2 = \sqrt{1 - c_0^2 - c_1^2}$ . Thus the futures-spot spread may depend on the risk factor  $W^2$  which cannot be hedged by trading in the futures market.

### 3 Mean-variance hedging

In this section we construct our optimal hedging strategies. After reviewing the key concepts and theorems from the mean-variance hedging literature in section 3.1, our main results are contained in sections 3.2 and 3.3 in which we solve the optimal hedging problems in our model, and apply these results to the fuel hedger. The proofs are postponed to the appendix.

We consider an agent with initial capital  $X_0$  who is exposed to a cumulative discounted cash flow  $H_t \in L^2(P)$  during  $[0, t]$ . In our applications,  $S_t$  is the spot price of a commodity (a crude oil derivate), and the agent is either

- a consumer who buys the commodity on the spot market at a constant rate per time; in this case  $H_t = -\int_0^t e^{-ru} S_u du$ , or
- a producer who sells the commodity on the spot market at a constant rate per time; in this case  $H_t = \int_0^t e^{-ru} S_u du$ .

The agent may trade in the futures market using a self-financing trading strategy  $\vartheta_t = (\vartheta_t^1, \dots, \vartheta_t^m)$ , where  $\vartheta_t^j$  is a predictable process denoting the number of futures contracts  $F(t, T_j)$  held at time  $t$ . To simplify the notation, we set  $\pi_t = e^{-rt}\vartheta_t$ . The discounted gains process from trading is then given by the stochastic integral

$$G_t(\pi) = \sum_{j=1}^m \int_0^t \vartheta_u^j e^{-ru} dF(u, T_j) = \sum_{j=1}^m \int_0^t \pi_u^j dF(u, T_j). \quad (3.1)$$

Hence the discounted value process  $X_t(\pi)$  of the agent's portfolio is

$$X_t(\pi) = X_0 + G_t(\pi) + H_t. \quad (3.2)$$

We fix a time horizon  $T \in [0, T^*]$ , and take a risk aversion parameter  $a > 0$ , a return level  $m \in \mathbb{R}$ , and a variance level  $v > 0$ . The agent's objective is to solve one of the following three optimization problems

$$U(a) = \sup_{\pi \in \mathcal{A}} \left( E[X_T(\pi)] - a \text{Var}[X_T(\pi)] \right), \quad (3.3)$$

$$B(m) = \inf_{\pi \in \mathcal{A}} \left\{ \text{Var}[X_T(\pi)] \mid E[X_T(\pi)] = m \right\}, \quad (3.4)$$

$$C(v) = \sup_{\pi \in \mathcal{A}} \left\{ E[X_T(\pi)] \mid \text{Var}[X_T(\pi)] = v \right\} \quad (3.5)$$

over all trading strategies  $\pi$  in a suitable set  $\mathcal{A}$  of admissible processes. We shall specify  $\mathcal{A}$  in Definition 3.1 below. The primal problems (3.3) – (3.5) are closely related to the classical quadratic hedging problem

$$A(\lambda) = \inf_{\pi \in \mathcal{A}} E \left[ (X_T(\pi) - \lambda)^2 \right] = \inf_{\pi \in \mathcal{A}} E \left[ (H(\lambda) - G_T(\pi))^2 \right] \quad (3.6)$$

with  $H(\lambda) = \lambda - X_0 - H_T$  for  $\lambda \in \mathbb{R}$ . Indeed, from the solution to  $A(\lambda)$  for all  $\lambda \in \mathbb{R}$ , one can immediately deduce the solution to (3.3) – (3.5), see Theorem 3.3 b) – d) below. In the following we solve (3.3) – (3.6) in the context of the Gaussian convenience yield model of section 2.4. For more background and general results on problem (3.6), we refer to Schweizer [21] and Černý and Kallsen [5].

### 3.1 Mean-variance hedging and variance-optimal martingale measure

In this section we review the key results from the literature on mean-variance hedging that we shall need here. We start by defining the market model and the set of admissible trading strategies. We resume the model setup of section 2.4. More precisely, we consider a futures market  $F(t, T_j)$  for  $j = 1, \dots, m$  on  $(\Omega, \mathcal{F}, P)$  satisfying (2.8) – (2.14). Set  $\vec{F}(t) = (F(t, T_j))_{j=1, \dots, m}$ . A trading strategy is called *simple* if it can be written as a linear combination of processes  $hI_{(\tau_1, \tau_2]}(t)$  where  $h$  is a bounded  $\mathcal{F}_{\tau_1}$ -measurable random variable and  $\tau_1 \leq \tau_2$  are stopping times such that  $\vec{F}(t \wedge \tau_2)$  is bounded. Loosely speaking, a strategy is admissible if it can be approximated by simple strategies. Recall the definition of the gains process in (3.1).

**Definition 3.1.** A predictable and  $\vec{F}$ -integrable process  $\pi$  is called an *admissible* trading strategy if there exist simple strategies  $\pi^{(n)}$ ,  $n \in \mathbb{N}$ , such that

$$\begin{aligned} G_t(\pi^{(n)}) &\rightarrow G_t(\pi) \quad \text{in probability for each } t \in [0, T] \\ G_T(\pi^{(n)}) &\rightarrow G_T(\pi) \quad \text{in } L^2 \end{aligned}$$

as  $n \rightarrow \infty$ . We define  $\mathcal{A} = \{\pi \mid \pi \text{ admissible}\}$ .

While this admissibility condition allows wealth processes that are unbounded from below, it excludes arbitrage opportunities. The set  $G_T(\mathcal{A})$  is closed in  $L^2(P)$  (see appendix A.2). This ensures that (3.6) has a solution. Next let

$$\mathcal{M} = \left\{ Q \approx P \mid \frac{dQ}{dP} \in L^2(P) \text{ and } \vec{F}(t) \text{ is a } Q\text{-local martingale} \right\}$$

denote the set of equivalent local martingale measures for  $\vec{F}(t)$  with square-integrable density. Let  $\mathcal{D} = \left\{ \frac{dQ}{dP} \mid Q \in \mathcal{M} \right\}$  be the set of densities of measures in  $\mathcal{M}$ . The following notion plays a central role in the solution of (3.6).

**Definition 3.2.** A measure  $Q$  in  $\mathcal{M}$  is called *variance-optimal* if it minimizes  $E[D^2]$  over all  $D \in \mathcal{D}$ .

The existence of variance-optimal measures is non-trivial since  $\mathcal{M}$  is not closed. We will show in Theorem 3.4 that under a suitable constraint on the model parameters,

$$\mathcal{M} \neq \emptyset. \tag{3.7}$$

Since  $\vec{F}$  is continuous, Theorem 1.3 in Delbaen and Schachermayer [7] then ensures that there exists a unique variance-optimal measure  $\tilde{P}$  in  $\mathcal{M}$ , called the *variance-optimal martingale measure (VOMM)* henceforth. Since  $D = \frac{d\tilde{P}}{dP} \in L^2(P)$ , we have  $\tilde{E}[D] = E[D^2] < \infty$  and we can define

$$\tilde{Z}_t := \tilde{E} \left[ \frac{d\tilde{P}}{dP} \mid \mathcal{F}_t \right], \quad t \in [0, T]. \tag{3.8}$$

By Lemma 2.2 in [7] it holds that

$$\tilde{Z}_t = \tilde{E} \left[ \frac{d\tilde{P}}{dP} \right] + \int_0^t \tilde{\zeta}_s \cdot d\vec{F}(s) \tag{3.9}$$

for some predictable process  $\tilde{\zeta}_t = (\tilde{\zeta}_t^1, \dots, \tilde{\zeta}_t^m)$ . The key result for the solution of (3.6) is



**Theorem 3.3.** *Assume (3.7). Write the Galtchouk-Kunita-Watanabe decomposition of  $H(\lambda)$  with respect to  $\vec{F}(t)$  under  $\tilde{P}$  as*

$$V_t(\lambda) = \tilde{E}[H(\lambda) | \mathcal{F}_t] = \tilde{E}[H(\lambda)] + \int_0^t \tilde{\xi}_s \cdot d\vec{F}(s) + L_t, \quad t \in [0, T]. \quad (3.10)$$

a) *The optimal control  $\pi_t = \pi_t(\lambda)$  in (3.6) is then given in feedback form by*

$$\pi_t(\lambda) = \tilde{\xi}_t - \frac{\tilde{\zeta}_t}{\tilde{Z}_t} \left( V_t(\lambda) - \int_0^t \pi_s(\lambda) \cdot d\vec{F}(s) \right), \quad t \in [0, T]. \quad (3.11)$$

b) *The optimal control in (3.3) is given by  $\pi_t(\lambda^*)$  with  $\lambda^* = \frac{\tilde{Z}_0}{2a} + X_0 + \tilde{E}[H_T]$ .*

c) *The optimal control in (3.4) is given by  $\pi_t(\lambda_m)$  with  $\lambda_m = \frac{\tilde{Z}_0 m - X_0 - \tilde{E}[H_T]}{\tilde{Z}_0 - 1}$ .*

d) *The optimal control in (3.5) is given by  $\pi_t(\lambda_{m(v)})$  with  $m(v) = X_0 + \tilde{E}[H_T] + \sqrt{(v - R)(\tilde{Z}_0 - 1)}$  for each  $v \geq R$ , where*

$$R = E \left[ \left( \tilde{Z}_T \int_0^T \frac{1}{\tilde{Z}_s} dL_s \right)^2 \right]. \quad (3.12)$$

The representation (3.10) is called the *Föllmer-Schweizer decomposition* of  $H(\lambda)$  with respect to the asset  $\vec{F}(t)$ . Several versions of Theorem 3.3 a) can be found in a number of papers including [18], [11] and [5]. A brief discussion of relations between these results and the proof can be found in appendix A.2.

**Remark.** A closer investigation of the proof of Theorem 3.3 b) shows that for  $a \rightarrow \infty$ , we obtain the unconstrained minimal achievable variance  $\text{Var}[X_T(\pi)]$  among all hedging strategies, and the corresponding expected terminal wealth is given by

$$E[X_T(\pi)] = X_0 + \tilde{E}[H_T],$$

see (A.5) and the following formula for  $m^*$ . ◇

## 3.2 The optimal hedging strategy for the Gaussian convenience yield model

In order to solve the feedback equation (3.11) for the optimal strategy  $\pi_t(\lambda)$  in Theorem 3.3, we need to find the VOMM  $\tilde{P}$ , compute the processes  $\tilde{Z}_t$  and  $\tilde{\zeta}_t$  in (3.8), (3.9), and find the Föllmer-Schweizer decomposition (3.10). To this end, first note that by Itô's representation theorem for local martingales under a Brownian filtration, every measure  $Q \approx P$  has a density process of the form

$$E \left[ \frac{dQ}{dP} \middle| \mathcal{F}_t \right] = \mathcal{E} \left( - \int \theta \cdot dW \right)_t := e^{-\int_0^t \theta_u \cdot dW_u - \frac{1}{2} \int_0^t \|\theta_u\|^2 du}, \quad t \in [0, T]$$

for some predictable process  $\theta$  such that  $E[\mathcal{E}(-\int \theta \cdot dW)_T] = 1$ . We denote this measure by  $Q = P^\theta$ . Girsanov's theorem then implies that if  $Q$  is an equivalent local martingale measure for  $\vec{F}(t)$ , it is of the form  $Q = P^{(\hat{\theta}_t^0, \hat{\theta}_t^1, \hat{\theta}_t^2)}$  for some predictable process  $\hat{\theta}_t^i$ .

We now establish (3.7) under a suitable constraint on the model parameters which can be expressed in terms of the ODE system given in Lemma A.2 in appendix A.3. We require that

$$\text{The ODE system (A.7) - (A.9) has a finite solution on } [0, T]. \quad (3.13)$$

If the model parameters satisfy

$$\beta + \gamma_1 \eta \geq \eta \sqrt{2(\gamma_0^2 + \gamma_1^2)},$$

then (3.13) is fulfilled for any  $T > 0$ . If this inequality does not hold, (3.13) is equivalent to an upper bound on the horizon  $T$  in terms of  $\beta, \eta, \gamma_0, \gamma_1$ , see Lemma A.3 in the appendix for this bound.

**Theorem 3.4.** *Assume (3.13). Let  $\hat{\theta}_t = (\hat{\theta}_t^0, \hat{\theta}_t^1, \hat{\theta}_t^2)$  with  $\hat{\theta}_t^2 = 0$ . Then the process*

$$Z_t := \mathcal{E}\left(-\int \hat{\theta} \cdot dW\right)_t, \quad t \in [0, T]$$

*is a square-integrable  $P$ -martingale. In particular,  $\frac{d\hat{P}}{dP} = Z_T$  defines a probability measure  $\hat{P} \in \mathcal{M}$ , and the process  $\hat{W}_t = W_t + \int_0^t \hat{\theta}_u du$  is a 3-dimensional  $\hat{P}$ -Brownian motion on  $[0, T]$ .*

Let  $\vec{M}(t)$  be the  $P$ -local martingale part in the canonical decomposition of the semimartingale  $\vec{F}(t)$ . It is easy to verify that the measure  $\hat{P}$  defined in Theorem 3.4 is the *minimal equivalent martingale measure (MEMM)* for  $\vec{F}(t)$ , introduced (in a different context) in Föllmer and Schweizer [10] as the unique equivalent local martingale measure  $\hat{P}$  such that any square integrable  $P$ -martingale orthogonal to  $\vec{M}(t)$  is also a  $\hat{P}$ -martingale. If the market price of risk  $\hat{\theta}_t$  is deterministic, it is well-known that the VOMM coincides with the MEMM,  $\tilde{P} = \hat{P}$ , see for instance Theorem 7 in Schweizer [19]. In contrast, if the market price of risk depends on exogenous stochastic factors, it is not difficult to show that in general  $\tilde{P} \neq \hat{P}$ , see Theorems 11 and 12 in Pham et al [16]. Unlike the MEMM, the VOMM is often difficult to construct explicitly in these situations.

In our model, the market price of risk  $\hat{\theta}_t$  is a stochastic process, and we find strong statistical evidence that  $\gamma_0$  is positive in our estimation results in section 4.1, so that  $\hat{\theta}_t$  is indeed non-deterministic in applications. However, the sub-market consisting of the traded assets  $\vec{F}(t)$  is *complete*. Indeed, the incompleteness of the model stems only from the presence of the additional asset  $S_t$  whose value cannot be replicated by trading in  $\vec{F}(t)$ . This allows us to show that the VOMM coincides with the MEMM.

**Theorem 3.5.** *Assume (3.13). Then*

a)  $\tilde{P} = \hat{P}$ .

b) *The process  $\tilde{Z}$  in (3.9) is given by*

$$\tilde{Z}_t = e^{x(T-t)+y(T-t)X_t+z(T-t)X_t^2} Z_t = \tilde{Z}_0 - \int_0^t \tilde{Z}_s \psi_s \cdot d\hat{W}_s, \quad t \in [0, T] \quad (3.14)$$

where  $x(\cdot), y(\cdot), z(\cdot)$  are the solution to the ODE system (A.7) - (A.9) in Lemmas A.2 and A.3 and

$$\psi_t = (\psi_t^0, \psi_t^1, \psi_t^2) = \left(\hat{\theta}_t^0, \hat{\theta}_t^1 - (\eta y(T-t) + 2\eta z(T-t)X_t), 0\right).$$

In particular, for any  $t < T_i < T_h$  we can write

$$\tilde{Z}_t = \tilde{Z}_0 + \int_0^t (\tilde{\zeta}_s^i, \tilde{\zeta}_s^h) \cdot d(F(t, T_i), F(t, T_h)) \quad (3.15)$$

with  $\tilde{\zeta}_t^i = \frac{v_h(t)}{w(t)F(t, T_i)} \tilde{Z}_t$  and  $\tilde{\zeta}_t^h = -\frac{v_i(t)}{w(t)F(t, T_h)} \tilde{Z}_t$ , where

$$\begin{aligned} v_k(t) &= \sigma \rho_0 \psi_t^1 - (\sigma \rho_1 - \eta e_1(\beta, T_k - t)) \psi_t^0, \quad k = i, h, \\ w(t) &= \sigma \rho_0 \eta (e_1(\beta, T_i - t) - e_1(\beta, T_h - t)). \end{aligned}$$

To apply these results to the fuel hedging problem, it remains to compute the Föllmer-Schweizer decomposition of  $H(\lambda)$ . We take the case of a commodity consumer, so with the notation introduced in the beginning of section 3 we have  $H_T = -\int_0^T e^{-ru} S_u du$  and

$$H(\lambda) = \lambda - X_0 - H_T = \lambda - X_0 + \int_0^T e^{-ru} S_u du. \quad (3.16)$$

For simplicity of notation, we assume that  $T$  is equal to the maturity date of a futures contract.

**Theorem 3.6.** *Assume (3.13). At each time  $t \in [0, T]$  set  $T_0 := t$ , let  $T_j = T_j(t)$  with  $j = 1, \dots, k = k(t)$  denote the maturity dates in  $(t, T]$ , and let  $T_i = T_i(t)$  and  $T_h = T_h(t)$  be two maturity dates  $t < T_i < T_h$ .*

a) *The discounted value process  $V_t(\lambda) = \tilde{E}[H(\lambda) \mid \mathcal{F}_t]$  is given by*

$$V_t(\lambda) = \lambda - X_0 + \int_0^t e^{-ru} S_u du + \sum_{j=1}^k e^{-T_j r} \int_{T_{j-1}}^{T_j} q_j(t, u) du F(t, T_j) \quad (3.17)$$

with

$$q_j(t, u) = e^{m_j(u-t) + n_1(u-t)Y_t + n_2(u-t)X_t},$$

and the functions  $m_j(\cdot), n_1(\cdot), n_2(\cdot)$  are given by (A.18) – (A.20) in the appendix.

b) *The Föllmer-Schweizer decomposition of  $H(\lambda)$  is given by*

$$\begin{aligned} dV_t(\lambda) &= \sum_{j=1}^k \left( e^{-T_j r} \int_{T_{j-1}}^{T_j} q_j(t, u) du \right) dF(t, T_j) \\ &\quad - \left( \frac{\sigma \rho_0 C_t^1 - (\sigma \rho_1 - \eta e_1(\beta, T_h - t)) C_t^0}{\sigma \rho_0 \eta (e_1(\beta, T_i - t) - e_1(\beta, T_h - t)) F(t, T_i)} \right) dF(t, T_i) \\ &\quad + \left( \frac{\sigma \rho_0 C_t^1 - (\sigma \rho_1 - \eta e_1(\beta, T_i - t)) C_t^0}{\sigma \rho_0 \eta (e_1(\beta, T_i - t) - e_1(\beta, T_h - t)) F(t, T_h)} \right) dF(t, T_h) \\ &\quad + dL_t \end{aligned} \quad (3.18)$$

with

$$\begin{aligned} C_t^0 &= \sum_{j=1}^k e^{-T_j r} F(t, T_j) \int_{T_{j-1}}^{T_j} q_j(t, u) c_0 \nu n_1(u - t) du, \\ C_t^1 &= \sum_{j=1}^k e^{-T_j r} F(t, T_j) \int_{T_{j-1}}^{T_j} q_j(t, u) (c_1 \nu n_1(u - t) + \eta n_2(u - t)) du, \end{aligned}$$

and a  $\tilde{P}$ -local martingale  $L_t$  orthogonal to  $\vec{F}(t)$ .

The last three theorems are proved in appendix A.3.

### 3.3 Application to the fuel hedging problem

We consider two agents: A kerosene consumer (such as an airline) who tries to optimally hedge her consumption costs, and a kerosene producer (such as an oil refinery) who tries to optimally hedge her production income. In both situations, the agent is assumed to be risk-averse, and employs a mean-variance optimal strategy on a given time interval.

For each agent, we compare the performance of a quadratic hedging strategy of type (3.4) or (3.5) with a competitor who is not engaged in hedging. Let us take the case of the commodity consumer (the case of the producer is analogous). Fix a time horizon  $T$ . A consumer with initial wealth  $X_0$  who does not hedge has value process  $X_t(0) = X_0 + H_t$  corresponding to the strategy  $X_t(\pi)$  in (3.2) with  $\pi = 0$ . The hedger now considers one of the following two Markowitz-type optimization problems.

- The hedger can aim to maximize the expectation for a given variance by considering

$$\text{Set } v = \text{Var}[X_T(0)] = E[H_T^2] - E[H_T]^2 \text{ and solve for the optimal } \pi \text{ for } C(v). \quad (3.19)$$

In this problem, the hedger aims to maximize the expectation of terminal wealth while achieving the same variance as the non-hedger.

- Alternatively, the hedger can aim to minimize the variance for a given expectation. The non-hedger's expected terminal wealth is given by  $E[X_T(0)] = X_0 + E[H_T]$ . On the other hand, by the remark after Theorem 3.3, the global minimal variance can be achieved with an expected terminal wealth given by  $X_0 + \tilde{E}[H_T]$ . Therefore it is reasonable to look at

$$\text{Set } m = X_0 + \max(E[H_T], \tilde{E}[H_T]) \text{ and solve for the optimal } \pi \text{ for } B(m). \quad (3.20)$$

In this problem, the hedger aims to minimize the variance of terminal wealth while achieving at least the same expectation as the non-hedger.

In order to solve (3.19) and (3.20), we need to find the expectation and variance of the cumulative consumption process  $H_T$ . For simplicity of notation, we assume that  $T$  is equal to the maturity date of a futures contract.

**Theorem 3.7.** *Recall  $H_T = -\int_0^T e^{-ru} S_u du$ . Set  $T_0 = 0$  and let  $T_j$  with  $j = 1, \dots, k$  denote the maturity dates in  $(0, T]$ . Then*

$$E[H_T] = -\sum_{j=1}^k e^{-T_j r} \int_{T_{j-1}}^{T_j} e^{\ell_j(u) + p(u)Y_0 + s_j(u)X_0} du F(0, T_j), \quad (3.21)$$

$$E[H_T^2] = \sum_{i=1}^k \sum_{j=1}^k e^{-(T_i + T_j)r} \int_{T_{i-1}}^{T_i} \left( \int_{T_{j-1}}^{T_j} q_{ij}(u, v) dv \right) du F(0, T_i) F(0, T_j) \quad (3.22)$$

where

$$q_{ij}(u, v) = I_{\{u \geq v\}} e^{\ell_i(u-v) + m_{ij}(v) + w(v)Y_0 + w_{ij}(v)X_0} + I_{\{v > u\}} e^{\ell_i(v-u) + m_{ij}(u) + w(u)Y_0 + w_{ij}(u)X_0}$$

and the functions  $p, s_j, \ell_j, w, w_{ij}, m_{ij}$  are given by (A.23) – (A.25) and (A.28) – (A.30) in the appendix.

The proof is given in appendix A.3. We now have all the ingredients to compute the optimal hedging strategies, and summarize the solutions to the optimization problems in the following

**Corollary 3.8.** *Set  $H_T = -\int_0^T e^{-ru} S_u du$  for the case of a commodity consumer (in the case of a commodity producer, we set  $H_T = \int_0^T e^{-ru} S_u du$ ).*

a) *The strategy which maximizes expectation of terminal wealth while achieving the same variance as the non-hedger is given by  $\pi_t(\lambda_{m(v)})$  in Theorem 3.3 with  $v = E[H_T^2] - E[H_T]^2$ .*

b) *The strategy which minimizes variance of terminal wealth while achieving at least the same expectation as the non-hedger is given by  $\pi_t(\lambda_m)$  in Theorem 3.3 with  $m = X_0 + \max(E[H_T], \tilde{E}[H_T])$ .*

*The quantities  $\tilde{E}[H_T]$ ,  $E[H_T]$ ,  $E[H_T^2]$ , and  $\tilde{\zeta}_t, \tilde{Z}_t, \tilde{\xi}_t, V_t$  in Theorem 3.3 can be computed by the formulas given in Theorems 3.5, 3.6 and 3.7.*

In section 4, we describe an implementation of these strategies, and test their performance numerically by running the strategies on market data.

## 4 Model calibration and numerical results

We first discuss the calibration of the model in section 4.1. To deal with the problem of estimating market price of risk parameters, we propose and numerically evaluate a calibration procedure based on growth optimal portfolio strategies. In section 4.2, we outline the implementation of the mean-variance hedging strategies, and present our numerical results on the performance of the hedging strategies in an out-of-sample test on market data.

## 4.1 Data and calibration

**Data.** Our data consists of daily NYMEX futures prices on WTI crude oil (\$ per barrel) with maturity dates of 1, 2, 3 and 4 months from February 1987 to December 2010 provided by the EIA, see [http://www.eia.gov/dnav/pet/pet\\_pri\\_fut\\_s1\\_d.htm](http://www.eia.gov/dnav/pet/pet_pri_fut_s1_d.htm), and daily spot prices of kerosene jet fuel (\$ per gallon) from April 1990 to December 2010, see [http://www.eia.gov/dnav/pet/pet\\_pri\\_spt\\_s1\\_d.htm](http://www.eia.gov/dnav/pet/pet_pri_spt_s1_d.htm). In the notation of the model in section 2, we have daily futures price and kerosene spot data at times  $t \in \mathcal{T} = \{t_1, \dots, t_n\}$ . For each fixed time  $t \in \mathcal{T}$  we have  $m = 4$  futures prices  $F(t, T_{tj})$  with maturity dates  $T_{tj} = T_j(t)$ ,  $j = 1, \dots, m$ , and kerosene spot prices  $S_t$ . Finally, we use a daily time series of average 1 month CD rates from 1987 to 2010 as provided by the Federal Reserve as a proxy for the short interest rate  $r_t$ , see <http://www.federalreserve.gov/releases/h15/data.htm>.

**Parameters.** The set of parameters for the futures market is  $(\sigma, \rho, \eta, \beta, \epsilon_0, \epsilon_1, \kappa, \beta_0, \gamma_0, \beta_1, \gamma_1)$  with the relations

$$\beta_1 = \sigma\rho - \frac{\epsilon_1}{\eta}, \quad (4.1)$$

$$\gamma_1 = \frac{\beta - \kappa}{\eta}, \quad (4.2)$$

see the proof of Theorem 2.2. A minimal set of parameters is therefore given by

$$(\sigma, \rho, \eta, \beta, \epsilon_0, \epsilon_1, \kappa, \beta_0, \gamma_0).$$

Here  $\sigma, \rho, \eta$  are determined by the covariance matrix of  $(dF_t, dX_t)$  in (2.8) and (2.10),  $\beta, \epsilon_0, \epsilon_1$  are determined by the shape of the futures price curve in (2.9), and  $\kappa, \beta_0, \gamma_0$  are determined by the drifts of  $F_t$  and  $X_t$ . Drift parameter estimates are typically afflicted with high parameter uncertainty. In the following we shall therefore separate the estimation problem for  $(\sigma, \rho, \eta, \beta, \epsilon_0, \epsilon_1)$  from the estimation of  $(\kappa, \beta_0, \gamma_0)$ . Finally, the set of parameters for the futures-spot spread process  $Y_t$  in (2.1), (2.15) is

$$(\phi, b, \nu, c_0, c_1).$$

**Estimation of volatility and curve shape parameters.** Let  $F(t, T) = F(t, T_{tj})$ ,  $t \in \mathcal{T}$ ,  $j = 1, \dots, m$  denote the futures price data. We will infer a time series  $(F, X) = (F_t, X_t)_{t \in \mathcal{T}}$  and parameters  $\Theta = (\eta, \beta, \epsilon_0, \epsilon_1)$  such that (2.9) is approximated in a least squares sense. More precisely, with the functions  $a$  and  $b$  defined in (2.12), (2.13), we compute  $X$  and  $\Theta$  by

$$\min_{X, \Theta} \sum_{t \in \mathcal{T}} \sum_{j=2}^m \left( a(T_{tj} - t) - a(T_{t1} - t) + (b(T_{tj} - t) - b(T_{t1} - t))X_t - (T_{tj} - T_{t1})r + \log F(t, T_{tj}) - \log F(t, T_{t1}) \right)^2 \quad (4.3)$$

under the constraints  $\frac{1}{n} \sum_{i=1}^n X_{t_i} = 0$  and  $\frac{252}{n-2} \sum_{i=2}^n (X_{t_i} - X_{t_{i-1}})^2 = \eta^2$ . We then set

$$F_t = F(t, T_{t1})e^{a(T_{t1}-t)+b(T_{t1}-t)X_t-(T_{t1}-t)r}, \quad t \in \mathcal{T},$$

which ensures that (2.9) holds exactly for  $F(t, T_{t1})$ . This is motivated by the fact that contract 1 is the most liquidly traded futures contract. Finally  $\sigma^2$  and  $\rho$  are computed as the annualized sample variance of  $\Delta \log F_t$  and sample correlation of  $\Delta X_t$  and  $\Delta \log F_t$ , respectively.

**Estimation of market price of risk parameters.** It is well-known that the estimation of the market price of risk, or equivalently, the drift parameters in asset price processes, is a difficult problem; estimators are usually afflicted with high uncertainty. In our model, we propose to estimate the market price of risk

parameters via a portfolio-based approach as follows. In the model (2.14) with market price of risk given by (2.11), we can choose any two fixed futures contracts to span the price risk in the futures market, that is, to replicate any other contract. We choose the futures contracts 1 and 3 for trading, and compute the self-financing trading strategy which is given by the *growth optimal portfolio* (GOP), see for instance Platen [17] for an overview on the theory of the GOP. Let  $\pi$  denote the position vector in the futures, and  $V^\pi$  the associated value process, of a self-financing trading strategy. It can then be shown (see section 5.3 in [17]) that the GOP strategy  $\pi^*$  maximizes the *long term growth rate*

$$g^\pi = \limsup_{T \rightarrow \infty} \frac{1}{T} \log V_T^\pi \quad (4.4)$$

over all self-financing trading strategies  $\pi$ . With given volatility and curve shape parameters  $\sigma, \rho, \eta, \beta$ , one can express the position vector  $\pi^*$  in the futures contracts for the GOP as a function of the associated value process  $V_t$ , the futures prices  $\vec{F}(t)$ , and the market price of risk  $\theta_t = (\theta_t^0, \theta_t^1)$ . Suppressing the dependence on  $V_t, \vec{F}(t)$ , and the volatility parameters, we write  $\pi^* = \pi^*(\theta)$ , and  $V = V^{\pi^*(\theta)}$  for the associated GOP value process. We now propose to estimate the parameters of the market price of risk  $\theta = \theta(\beta_0, \gamma_0, \beta_1, \gamma_1)$  by maximizing  $g^{\pi^*(\theta)}$  over all parameters values  $(\beta_0, \gamma_0, \beta_1, \gamma_1)$ . As a proxy for the (non-observable) long term growth rate  $g^\pi$  in (4.4), we assume that the lim sup is equal to the limit, and estimate it by an average of  $\frac{1}{T} \log V_T^\pi$  over all sufficiently large  $T$ . We then obtain

$$\hat{\theta} = \operatorname{argmax}_{\theta} \frac{1}{n-k+1} \sum_{i=k}^n \frac{1}{t_i} \log V_{t_i}^{\pi^*(\theta)} \quad (4.5)$$

as our estimator for the market price of risk.

We implement two versions of this estimation method. In the first method, the maximization in (4.5) is performed over all parameters  $\beta_0, \gamma_0, \beta_1, \gamma_1$  in a single step, using  $k = n$ . The resulting estimates for  $\beta_1$  and  $\gamma_1$  exhibit a rather unstable behavior over time, including regime changes in times of market stress, and the associated GOP value process has large jumps. We then compute  $\beta_1$  and  $\gamma_1$  via the formulas (4.1), (4.2), where  $\kappa$  is estimated via linear regression on the time series  $X_t$  with  $t$  starting in year 1991. A comparison suggests that the two alternative estimation procedures seem to converge for sufficiently large sample sizes (the parameter estimates are plotted in Figure 4 below), but again estimates are rather unstable over time. These leads us to conjecture that there does not exist a long-term trend in the  $W_t^1$  risk factor of the futures price dynamics, that is, in fact  $\hat{\theta}_t^1 = 0$ . To test this conjecture, for each parameter  $\beta_1$  and  $\gamma_1$  and each estimation procedure (GOP-based and via (4.1), (4.2)), we fit the time series of the parameter estimates to an ARMA(1,1) model, and test the null hypothesis that the ARMA(1,1) intercept is zero. For the GOP-based estimates, we find that the null hypothesis cannot be rejected for  $\beta_1$  on the 5% level, but it is rejected for  $\gamma_1$ . Using the formulas (4.1), (4.2), we find that the null hypothesis cannot be rejected on the 5% level for both parameters, supporting the conjecture  $\hat{\theta}_t^1 = 0$ .

Motivated by these findings, we implement a second version of the estimation method in which we set  $\beta_1 = \gamma_1 = 0$ , and perform the maximization in (4.5) only over  $\beta_0, \gamma_0$ . We find the estimates for  $\beta_0, \gamma_0$  to be considerably more stable, and a similar test as above clearly rejects the hypotheses that  $\beta_0, \gamma_0$  are zero (the parameter estimates are plotted in Figure 5 below; here we chose  $k < n$  to further increase parameter stability). We implement the corresponding growth optimal portfolio strategy in an out-of-sample test by recalibrating the model at each time  $t$  based on information available at  $t$ . The discounted wealth process  $V_t^{\pi^*(\theta)} e^{-\int_0^t r_u du}$  for  $V_0^{\pi^*(\theta)} = 1$  is plotted in Figure 2. It shows a clear long-term upward trend.

**Estimation of spread process.** The process  $Y_t$  in (2.15) follows an Ornstein-Uhlenbeck process. We estimate  $\phi, b, \nu$  by maximum likelihood (or equivalently, by linear regression), and  $c_0, c_1$  via the sample

correlations of  $\Delta Y_t$ ,  $\Delta X_t$  and  $\Delta Y_t$ ,  $\Delta F_t$ .

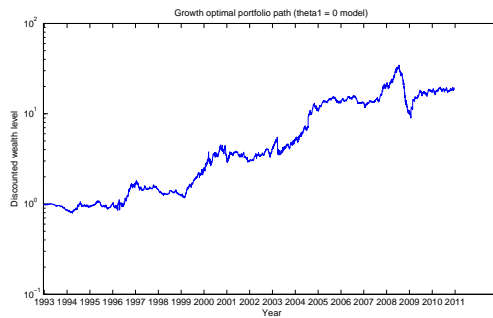


Figure 2: Discounted growth optimal portfolio process for  $(\beta_0, \gamma_0)$ -optimization. Note logarithmic scale.

**Calibration results.** In order to assess the potential out-of-sample performance of our hedging strategies, we are interested in the stability of the model parameters over time. To this end, we employ a recalibration procedure as follows. For each parameter  $x$ , we compute a time series  $x_i$  of estimates, where the estimator  $x_i$  is based on the data at times  $t_1, \dots, t_i$ . The estimates start at day  $i = 1511$ , that is we use an initial window of about 6 years (covering the period from 1987 to 1993). We found the parameter estimates to be quite unreliable for smaller initial windows.

*a) Volatility and curve shape parameters.* Our results are plotted in Figure 3. It can be seen that the volatility and curve shape parameters  $\sigma, \rho, \eta, \beta$  are very stable over time, whereas the curve shape parameters  $\epsilon_0, \epsilon_1$  are considerably less stable.

*b) Market price of risk parameters.* For the market price of risk, we first compute the parameters  $(\beta_0, \gamma_0, \beta_1, \gamma_1)$  via the GOP-based estimation procedure in (4.5) and in addition  $\beta_1$  and  $\gamma_1$  via (4.1), (4.2). The results are plotted in Figure 4. We next set  $\hat{\theta}_t^1 = 0$  and estimate  $(\beta_0, \gamma_0)$  again via the GOP-based estimation procedure in (4.5). The corresponding estimates for  $(\beta_0, \gamma_0)$  are plotted in Figure 5, and can be seen to be quite stable over time.

*c) Spread process parameters.* Here our data and the initial calibration window starts in 1990. Our results are plotted in Figure 6.

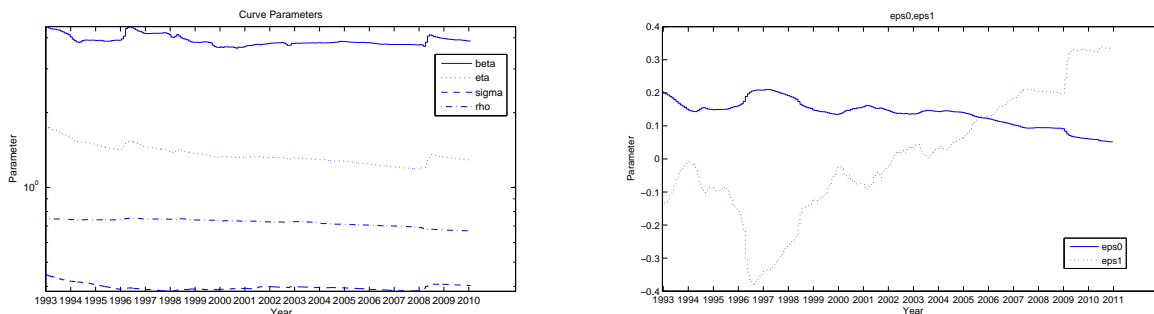


Figure 3: Estimates for futures volatility and curve shape parameters, based on data from 1987 to indicated year.

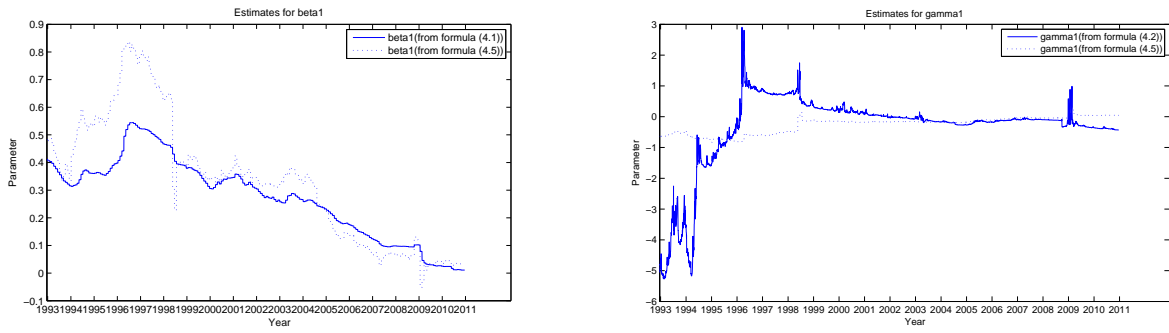


Figure 4: Estimates for  $\beta_1$  and  $\gamma_1$  based on formulas (4.1), (4.2) and GOP-approach (4.5), using data from 1987 to indicated year.

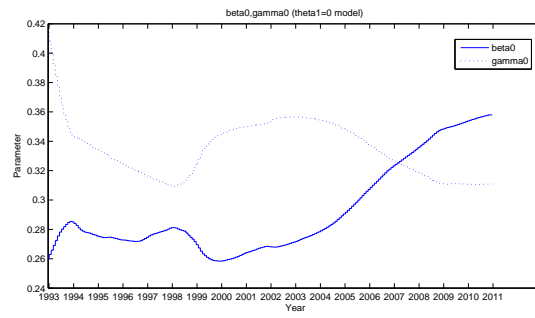


Figure 5: GOP-based estimates for  $\beta_0$  and  $\gamma_0$ , using data from 1987 to indicated year.

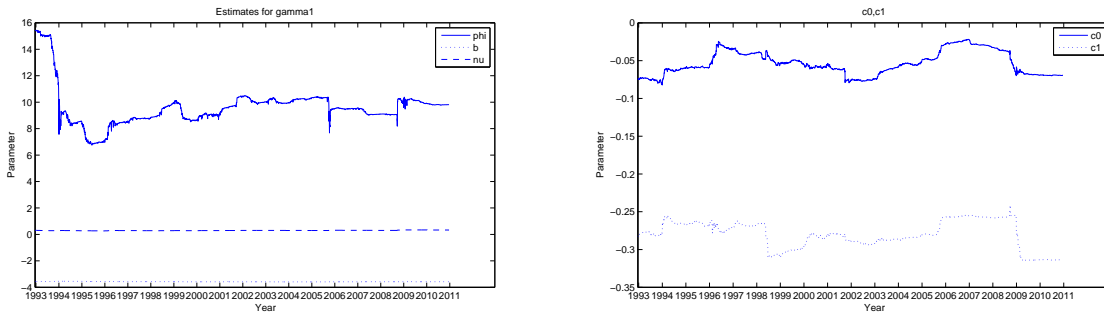


Figure 6: Estimates for spread process parameters, based on data from 1990 to indicated year.



## 4.2 Implementation of the quadratic hedging strategies and numerical results

We perform an out-of-sample test for the quadratic hedging strategies as follows. We use the first 6 years of daily data for initial calibration of the model as described in section 4.1, and after the initial period we recalibrate the model at each day  $i$  using all data up to day  $i$ . The out-of-sample test period after the initial calibration period is divided into intervals of equal length  $T = 3$  months, and we implement the optimal quadratic hedging strategies in Corollary 3.8 separately on each of these intervals, using daily portfolio re-balancing in our trading strategies. In total we have 72 intervals in the out-of-sample test period, covering the time from 1993 to 2010. We assume that there are no transaction costs for trading in the futures market. The feedback equation for the optimal strategy position  $\pi_t(\lambda)$  is implemented via the explicit discretization scheme

$$\pi_{t_i}(\lambda) = \tilde{\xi}_{t_i} - \frac{\tilde{\zeta}_{t_i}}{\tilde{Z}_{t_i}} \left( V_{t_i}(\lambda) - \sum_{j \leq i-1} \pi_{t_j}(\lambda) \cdot (\bar{F}(t_{j+1}) - \bar{F}(t_j)) \right).$$

Except for the value of  $R$  in (3.12) (which enters into the formula for  $\lambda = \lambda_{m(v)}$  in the maximization of expectation problem), all terms in this equation are given via explicit formulas. We approximate  $R$  via an analytic formula which is based on approximating the market price of risk by its initial value and using Theorem 2 in [2]. (We conjecture that an analytical formula could also be obtained for the exact value of  $R$  in our model.) To test the accuracy of the approximation formula, we also compute  $R$  at the beginning of each quadratic hedging interval from (A.4) by estimating  $B(m) = \text{Var}[X_T(\pi)]$  with  $\pi = \pi(\lambda_m)$  for  $m = X_0 + \tilde{E}[H_T]$  via a Monte Carlo simulation. We find that the approximation formula underestimates  $R$  on average by about a factor 2. We then analyze the effect of this approximation error to  $\lambda_{m(v)}$  (which determines the optimal strategy), and find that the resulting relative error in  $\lambda_{m(v)}$  is very small (less than 1% in all but three intervals in which it reaches at most about 10%).

We implement the optimal hedging strategy for the expectation maximization problem (3.19) in Corollary 3.8 a) and the optimal hedging strategy for the variance minimization problem (3.20) in Corollary 3.8 b). We assume a constant consumption or production rate of one gallon of jet fuel per year. All implementations use the  $\hat{\theta}_t^1 = 0$  model, see section 4.1 for a discussion of statistical support for this choice.

**Remark.** The  $\hat{\theta}_t^1$  component of the MPR determines trends in the futures term structure factor  $X_t$ , while  $\hat{\theta}_t^0$  determines trends in the crude spot price that are independent of  $X_t$ . Aside from statistical considerations, one might also argue that the risk manager of a non-financial corporation would prefer to concentrate on “trading trends in the spot price” rather than “betting on the futures term structure”.

The results for the jet fuel consumer are summarized in Figures 7 and 8. For the expectation maximization problem, the first diagram in Figure 7 shows the total cash flow from jet fuel consumption in each period for both the hedger (green) and the non-hedger (red). In a few hedging periods, the total cash flow is positive, meaning that the hedging strategy generates gains which exceed the cost of fuel consumption in those periods. On the other hand, in the 2008 oil market crash, the hedging strategy generates considerable losses in two periods (the columns are cut off in the diagram and have values  $-3$  and  $-4.5$ ). In the second diagram, we compare the cumulative wealth from consumption and trading for the hedger and the non-hedger since 1993 at two time points before and after the 2008 crash. Finally, the third diagram shows the cumulative wealth processes for the hedger and the non-hedger on a daily basis.

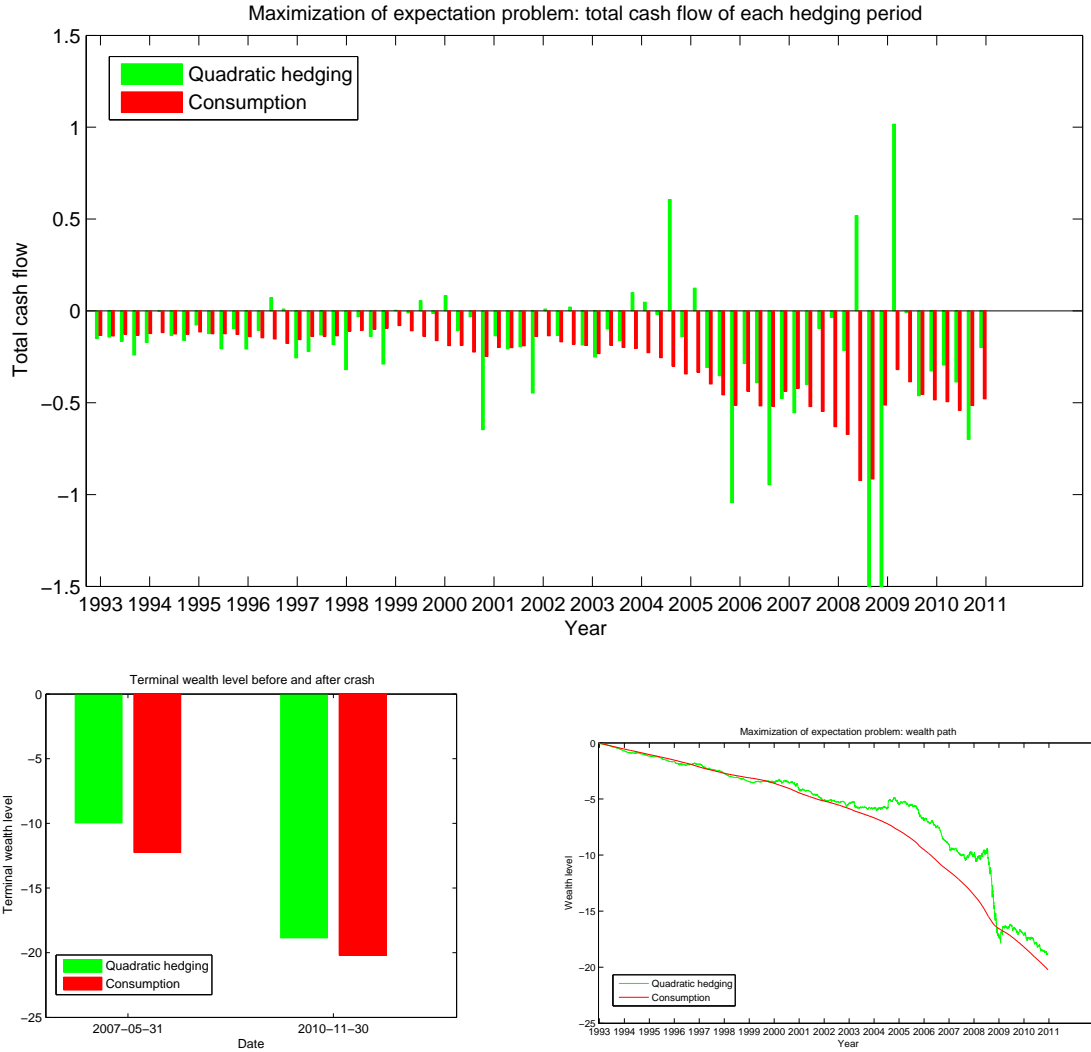


Figure 7: Maximization of expectation for jet fuel consumer.

For the variance minimization problem, the first diagram in Figure 8 again compares the total cash flow in each period for the hedger (green) and the non-hedger (red). The second diagram shows the deviations of the end-of-period realized cash flow (or terminal wealth) from the *expected terminal wealth*, again separately for the hedger and the non-hedger. The realized standard deviations from expected terminal wealth over all 72 periods for the hedger and the non-hedger are reported in Table 1.



Figure 8: Minimization of variance for jet fuel consumer.

Finally, the results for the jet fuel producer are summarized in the same way in Figure 9 for the expectation maximization problem (the outliers in the first diagram are at  $-2.5$  and  $2.5$ ) and in Figure 10 for the variance minimization problem. For the variance minimization problem, the realized standard deviations from expected terminal wealth over all 72 periods for the hedger and the non-hedger are reported in Table 1.

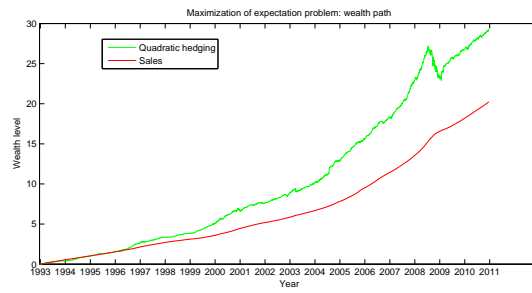
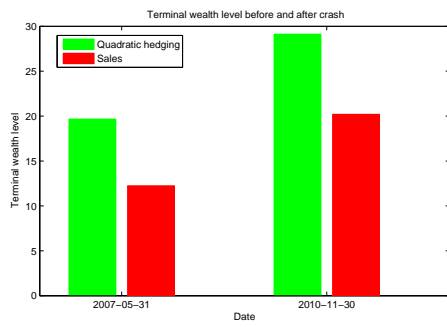
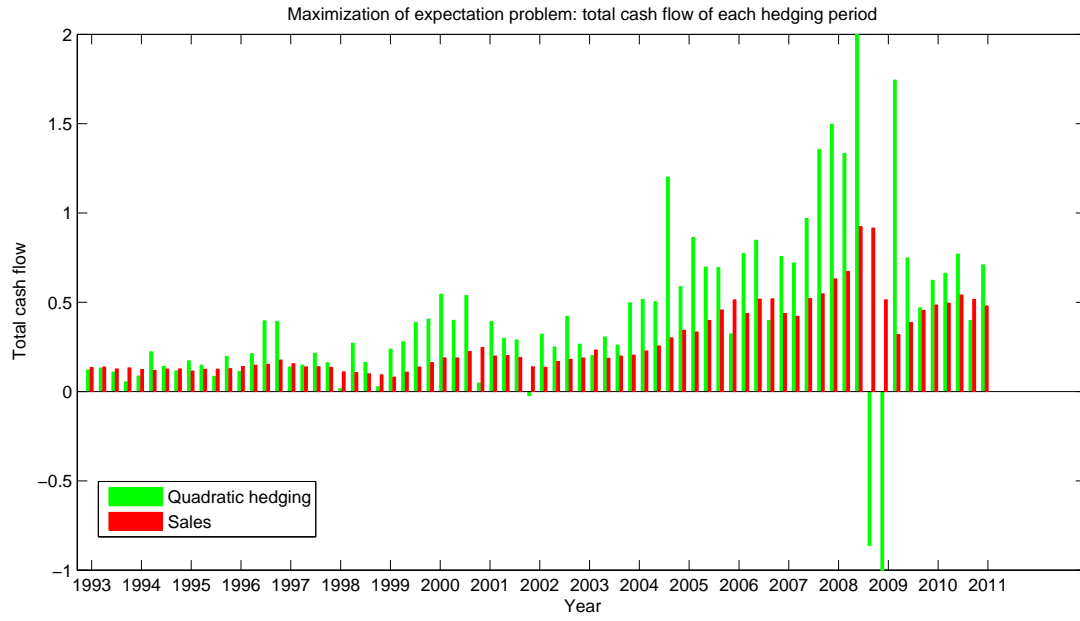


Figure 9: Maximization of expectation for jet fuel producer.

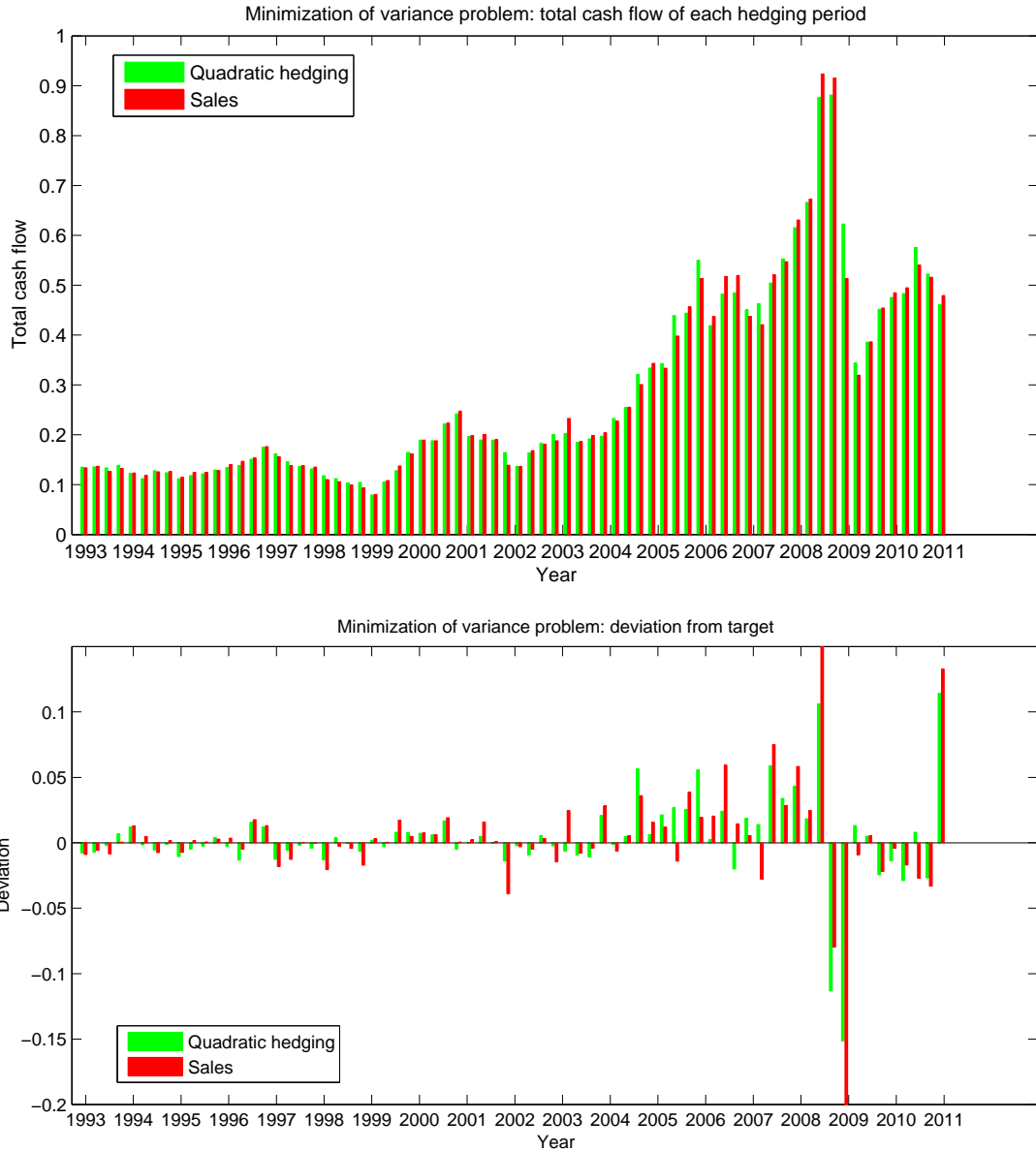


Figure 10: Minimization of variance for jet fuel producer.

By the duality results in Theorem 3.3, each optimal strategy in problem (3.19) and (3.20) corresponds to an optimal strategy for the mean-variance tradeoff problem (3.3) with a specific choice of the risk aversion coefficient  $a$ . The numerical value of  $a$  depends on the problem and on the state of the market at the beginning of the hedging period, and can easily be deduced from the formulas in Theorem 3.3. For both the expectation maximization problem and the variance minimization problem, we compute the median value of  $a$  across all 72 hedging periods. The results are reported in Table 1. Clearly, the risk aversion coefficient  $a$  is higher for the variance minimization problem than for the expectation maximization. In some sense, the two problems correspond to two extreme cases of improving the non-hedger's risk versus return profile. Therefore any value between the expectation maximization and the variance minimization risk coefficient could be considered an economically reasonable choice for  $a$ . The actual choice of  $a$

depends on the agent’s subjective risk profile.

	median risk aversion $a$		standard deviations from target	
	max expectation	min variance	min variance hedging	no hedging
consumer	0.40	$\infty$	0.025	0.045
producer	0.40	6.6	0.034	0.045

Table 1: Summary statistics on risk aversion  $a$  and results for variance minimization problem.

## 5 Comments and conclusion

In this paper, we have studied optimal hedging strategies for an economic agent who is exposed to commodity price risk. We consider a market in which this risk can only be partially hedged by using futures contracts on a proxy commodity. We formulate and solve three related mean-variance optimal dynamic hedging problems in a continuous-time model. This leads to interesting problems in stochastic analysis due to the appearance of a *stochastic market price of risk process* in the underlying futures market model. We find analytical solutions for the optimal hedging problems in our market model.

The explicit solutions allow a fast algorithm for numerical evaluations. We apply our model to the case of hedging kerosene jet fuel via crude oil futures contracts. We calibrate our model using a portfolio-based estimation procedure for the market price of risk process. We find strong statistical evidence that the market price of risk (the component  $\hat{\theta}_t^0$ ) is indeed non-deterministic in the crude oil futures market. We then implement the optimal strategies and evaluate them in an out-of-sample test on a 18 year test period. Our results show that quadratic hedging improves the long-term performance of both a kerosene consumer and a kerosene producer. Moreover, our approach allows us to analyze the choice of the risk-aversion coefficient in the mean-variance optimal investment problem in a quantitative way.

The commodity market crash in the financial crisis of 2008 has a significant negative impact on the performance of the optimal hedging strategies, with the effect being stronger for the commodity consumer than for the producer due to the consumer’s long positions during the crash. The constant parameter assumption of our model is probably an important reason for the weak performance during this extreme market event. We conjecture that in a stochastic volatility model, an agent using mean-variance type optimality criteria would take less risky positions during times of market stress, reducing the impact of financial hedging in comparison to a constant parameter model. We are working to extend our model and hedging approach to stochastic volatility models in future research.

## A Appendix

All proofs are given in this appendix. Sections A.1 and A.2 review some key results from the literature on affine models and on quadratic hedging that we use in this paper. Section A.3 contains the proofs of our main results.

### A.1 Proof of Theorem 2.2

Assumption (2.7) implies that

$$\epsilon(t, T) = A(T - t) + B(T - t)X_t$$

with  $A(\tau) = a'(\tau)$  and  $B(\tau) = b'(\tau)$ , and thus by (2.8)

$$\begin{aligned} d\epsilon(t, T) &= -A'(T-t)dt - B'(T-t)X_t dt + B(T-t)dX_t \\ &= \left( -A'(T-t) - (B'(T-t) + \kappa B(T-t))X_t \right) dt + B(T-t)\eta dW_t^1 \end{aligned}$$

for all  $T > t$ . Comparing with (2.4) we obtain  $\sigma_\epsilon(t, T) = B(T-t)\eta$ , and then (2.5) and (2.11) yield

$$-A'(T-t) - (B'(T-t) + \kappa B(T-t))X_t = B(T-t)\eta \left( \beta_1 + \gamma_1 X_t + \int_t^T B(s-t)\eta ds - \sigma\rho \right).$$

By separating this into a sum of deterministic terms and linear terms in  $X_t$ , we obtain

$$\begin{aligned} B'(T-t) &= -(\kappa + \gamma_1\eta)B(T-t), \\ A'(T-t) &= B(T-t)\eta \left( \sigma\rho - \beta_1 - \int_t^T B(s-t)\eta ds \right). \end{aligned}$$

Since  $B(0) = b'(0) = 1$  we obtain  $B(T-t) = e^{-\beta(T-t)}$  with  $\beta = \kappa + \gamma_1\eta$  and

$$\begin{aligned} A(T-t) &= A(0) + \eta(\sigma\rho - \beta_1) \int_t^T B(s-t)ds - \frac{\eta^2}{2} \left( \int_t^T B(s-t)ds \right)^2 \\ &= \epsilon_0 + \epsilon_1 e_1(\beta, T-t) - \frac{\eta^2}{2} e_1(\beta, T-t)^2 \end{aligned}$$

with  $\epsilon_0 = A(0)$  and  $\epsilon_1 = \eta(\sigma\rho - \beta_1)$ . Integrating  $A$  and  $B$  then yields (2.12) and (2.13). Finally (2.14) follows by applying Itô's formula to (2.2).  $\blacksquare$

## A.2 General results on the mean-variance hedging problem

In this appendix we collect some technical background on the material in section 3.1 and provide references to the relevant literature on the general mean-variance hedging problem. For a continuous semimartingale  $X$  on  $[0, T]$  with canonical decomposition  $X_t = X_0 + M_t + A_t$ , we write  $X \in \mathcal{S}^2(P)$  if

$$\|X\|^2 := E \left[ X_0^2 + \langle M, M \rangle_T + \left( \int_0^T |dA_s| \right)^2 \right] < \infty.$$

Recall the set  $\mathcal{A}$  of admissible strategies in Definition 3.1. Various authors have worked with a different set  $\Theta$  of admissible strategies. Define

$$\Theta = \left\{ \pi \text{ predictable and } X\text{-integrable} \mid G(\pi) \in \mathcal{S}^2(P) \right\}.$$

**Theorem A.1.** *a) We have*

$$\mathcal{A} = \left\{ \pi \text{ predictable and } \vec{F}\text{-integrable} \mid G_T(\pi) \in L^2(P) \text{ and } G_t(\pi) \text{ is a } Q\text{-martingale for each } Q \in \mathcal{M} \right\}.$$

*b)  $G_T(\mathcal{A})$  is the closure of  $G_T(\Theta)$  in  $L^2(P)$ .*

Part a) is Theorem 2.8 in Černý and Kallsen [6], where the inclusion  $\supseteq$  follows from Theorems 1.2 and 2.2 in Delbaen and Schachermayer [8], and part b) is Corollary 2.9 part 1 in Černý and Kallsen [5]. In particular, a) says that the set of admissible strategies  $\mathcal{A}$  coincides with the set of strategies used in Gourieroux, Laurent and Pham [11]. Part b) in particular says that  $G_T(\mathcal{A})$  is closed in  $L^2(P)$ .  $\diamond$

**On the proof of Theorem 3.3.** a) This can be deduced as a special case of Theorem 4.10 in Černý and Kallsen [5], who establish the structure of the solution to the mean-variance hedging problem for a general semimartingale. This generality is not necessary for our setup, and simpler proofs can be found in the literature when asset prices are continuous. Theorem 3.3 is a direct restatement of Theorem 6 in Rheinländer and Schweizer [18]. However, the latter result cannot directly be applied since the authors work with the set  $\Theta$  of admissible strategies, which imposes a stronger integrability condition on the gains process. This leads to a condition on the asset price process which is necessary to ensure that  $G_T(\Theta)$  is closed in  $L^2(P)$ . For our market model, this condition is not satisfied (indeed, it is easy to see that  $Z_t$  does not satisfy the reverse Hölder inequality  $R_2(P)$  in Theorem 2 of [18]). The set  $\mathcal{A}$  of admissible trading strategies is the same as in Gourieroux, Laurent and Pham [11], whose setup applies to a general continuous semimartingale model. The main idea of [11] is to use  $\tilde{Z}_t$  is a numeraire (which is possible due to (3.9)). This reduces the mean-variance hedging problem to finding the Galtchouk-Kunita-Watanabe decomposition of the discounted cash flow under the risk-neutral measure  $\tilde{R}$  associated to numeraire  $\tilde{Z}_t$ . The optimal solution in Theorem 5.1 of [11] is expressed in the coordinates of the new numeraire, and therefore takes a form different from the feedback formula (3.11). It is shown in Propositions 8 and 10 of [18] that the formulas for the optimal solution in [11] and [18] are equivalent, which proves (3.11).

b) – d) The minimal value in (3.6) is given by

$$A(\lambda) = \frac{(\lambda - X_0 - \tilde{E}[H_T])^2}{\tilde{Z}_0} + R \quad (\text{A.1})$$

with  $R = E\left[\left(\tilde{Z}_T \int_0^T \frac{1}{\tilde{Z}_s} dL_s\right)^2\right]$ . This follows from rewriting (5.3) in [11] under  $\tilde{P}$  and  $P$ , and using the relation (4.13) in [18] between the Galtchouk-Kunita-Watanabe decompositions of the discounted cashflows under the measures  $\tilde{P}$  and  $\tilde{R}$ , respectively. Next for each  $m \in \mathbb{R}$  define

$$B(m) = \inf_{\pi \in \mathcal{A}} \left\{ \text{Var} [X_T(\pi)] \mid E[X_T(\pi)] = m \right\}. \quad (\text{A.2})$$

The same proof as for Proposition 6.6.5 in [15] shows that

$$B(m) = \sup_{\lambda \in \mathbb{R}} \left( A(\lambda) - (m - \lambda)^2 \right), \quad (\text{A.3})$$

and if  $\lambda_m$  is a maximizer in (A.3), the process  $\pi_t(\lambda_m)$  in (3.11) is an optimal control for  $B(m)$  in (A.2). Using (A.1), straightforward calculations yield the maximizer  $\lambda_m = \frac{\tilde{Z}_0 m - X_0 - \tilde{E}[H_T]}{\tilde{Z}_0 - 1}$  and the value

$$B(m) = \frac{(X_0 + \tilde{E}[H_T] - m)^2}{\tilde{Z}_0 - 1} + R. \quad (\text{A.4})$$

This yields c). For given  $v \geq R$ , d) follows from setting  $B(m) = v$  and solving the quadratic equation for its largest root. Finally, by definition of  $U(a)$  and  $B(m)$  we have

$$U(a) = \sup_{m \in \mathbb{R}} \left( m - aB(m) \right), \quad (\text{A.5})$$

and by (A.4) the maximum in (A.5) is attained at  $m^* = \frac{\tilde{Z}_0 - 1}{2a} + X_0 + \tilde{E}[H_T]$ . Moreover the optimal solution to (3.3) is now given by the optimal control to  $B(m^*)$ , which by the above observation is  $\pi_t(\lambda_{m^*})$ . Combining the formulas for  $\lambda_m$  and  $m^*$  yields b).  $\blacksquare$



### A.3 Proofs of Theorems 3.4 – 3.7

We resume the setup and definitions in Theorem 3.4. We start with

**Lemma A.2.** *Let  $x(\cdot), y(\cdot), z(\cdot)$  be differentiable functions with  $x(0) = y(0) = z(0) = 0$ . The process*

$$Y_t := e^{x(T-t)+y(T-t)X_t+z(T-t)X_t^2} Z_t^2 \quad (\text{A.6})$$

for  $t \in [0, T]$  is a  $P$ -local martingale if and only if  $x(\cdot), y(\cdot), z(\cdot)$  are a finite solution to the ODE system

$$z'(\tau) = \gamma_0^2 + \gamma_1^2 - 2(\beta + \gamma_1\eta)z(\tau) + 2\eta^2 z(\tau)^2, \quad (\text{A.7})$$

$$y'(\tau) = 2(\beta_0\gamma_0 + \beta_1\gamma_1) - (\beta + \gamma_1\eta)y(\tau) - 4\beta_1\eta z(\tau) + 2\eta^2 y(\tau)z(\tau), \quad (\text{A.8})$$

$$x'(\tau) = \beta_0^2 + \beta_1^2 - 2\beta_1\eta y(\tau) + \frac{1}{2}\eta^2 y(\tau)^2 + \eta^2 z(\tau) \quad (\text{A.9})$$

on  $[0, T]$ . In this case the process  $\hat{Z}_t := \frac{Y_t}{Z_t}$  satisfies  $\hat{Z}_T = Z_T$  and

$$\hat{Z}_t = \hat{Z}_0 - \int_0^t \hat{Z}_s \psi_s \cdot d\hat{W}_s, \quad t \in [0, T], \quad (\text{A.10})$$

where  $\hat{W}_t = W_t + \int_0^t \hat{\theta}_u du$  and  $\psi_t = \left( \hat{\theta}_t^0, \hat{\theta}_t^1 - (\eta y(T-t) + 2\eta z(T-t)X_t), 0 \right)$ .

*Proof.* Recall from (2.11) that  $\hat{\theta}_t = (\beta_0 + \gamma_0 X_t, \beta_1 + \gamma_1 X_t, 0)$ . From (2.8) we have

$$\begin{aligned} dX_t &= -\kappa X_t dt + \eta dW_t^1, \\ dX_t^2 &= (\eta^2 - 2\kappa X_t^2) dt + 2X_t \eta dW_t^1, \\ dZ_t^2 &= Z_t^2 (\|\hat{\theta}_t\|^2 dt - 2\hat{\theta}_t \cdot dW_t). \end{aligned}$$

Applying Itô's formula to (A.6),

$$\begin{aligned} dY_t &= Y_t \left( -2\hat{\theta}_t^0 dW_t^0 + (\eta y(T-t) + 2\eta z(T-t)X_t - 2\hat{\theta}_t^1) dW_t^1 \right) \\ &\quad + Y_t \left( -x'(T-t) - y'(T-t)X_t - z'(T-t)X_t^2 - y(T-t)\kappa X_t + z(T-t)(\eta^2 - 2\kappa X_t^2) + \|\hat{\theta}_t\|^2 \right. \\ &\quad \left. - 2\hat{\theta}_t^1 (\eta y(T-t) + 2\eta z(T-t)X_t) + \frac{1}{2} (\eta y(T-t) + 2\eta z(T-t)X_t)^2 \right) dt. \end{aligned} \quad (\text{A.11})$$

Using  $\kappa = \beta - \gamma_1\eta$  and writing the drift in (A.11) as a quadratic function in  $X_t$  with deterministic coefficients, we obtain that the drift vanishes (that is,  $Y_t$  is a  $P$ -local martingale) if and only if (A.7) - (A.9) hold true. Finally, we note that Itô's formula, (A.11) and  $dZ_t = -Z_t \hat{\theta}_t \cdot dW_t$  imply

$$d\hat{Z}_t = \hat{Z}_t \left( -\hat{\theta}_t^0 d\hat{W}_t^0 + (\eta y(T-t) + 2\eta z(T-t)X_t - \hat{\theta}_t^1) d\hat{W}_t^1 \right)$$

which gives (A.10). ■

The solution of the Ricatti equation system (A.7) - (A.9) can be expressed in closed form.

**Lemma A.3.** *For constant coefficients  $a, b, c, f, h, k \in \mathbb{R}$  with  $a, c > 0$ , define  $d = \sqrt{b^2 - 4ac}$  and  $g = \frac{d+b}{d-b}$ . Then the ODE system*

$$\begin{aligned} z'(\tau) &= a + b z(\tau) + c z(\tau)^2, & z(0) &= 0, \\ y'(\tau) &= f + \frac{b}{2} y(\tau) + h z(\tau) + c y(\tau)z(\tau), & y(0) &= 0, \\ x'(\tau) &= k + \frac{b}{2} y(\tau) + \frac{1}{4} c y(\tau)^2 + \frac{1}{2} c z(\tau), & x(0) &= 0 \end{aligned}$$

has the solution

$$\begin{aligned} z(\tau) &= \frac{2a}{d-b} \frac{1 - e^{-d\tau}}{1 + ge^{-d\tau}}, \\ y(\tau) &= \frac{2}{(d-b)d} \frac{1}{1 + ge^{-d\tau}} \left( f(d-b) + 2ha - (4ha + f(1-g)(d-b))e^{-\frac{1}{2}d\tau} + (2ha - fg(d-b))e^{-d\tau} \right), \\ x(\tau) &= c_1 + c_2\tau + \frac{1}{(d-b)^2 d^4} \frac{1}{1 + ge^{-d\tau}} \left( c_3 + c_4 e^{-\frac{1}{2}d\tau} + c_5 e^{-d\tau} \right) + \frac{1}{2} \log \left( \frac{1+g}{1+ge^{-d\tau}} \right) \end{aligned}$$

with

$$\begin{aligned} c_1 &= -\frac{c(-3bf+df+6ah)(f(b+d)-2ah)}{2d^4(b+d)} \\ c_2 &= -\frac{b+d}{4} + k + \frac{cf^2 - bhf + ah^2}{d^2} \\ c_3 &= 4cf^2b^3 + (b+d)(-7cf^2 + 2ah^2)b^2 + (2adh^2 + cf(6ah + 5fd))b + 2ac(3cf^2 + h(ah - 3fd)) \\ c_4 &= 4d(d-b)(bf - 2ah)(bh - 2cf) \\ c_5 &= 2((cf^2 + ah^2)b^3 - (adh^2 + cf(df + 6ah))b^2 - ac(cf^2 + h(ah - 10df))b - ac(7adh^2 + cf(5df - 12ah))). \end{aligned}$$

The above formulas are to be understood as their analytic continuation if  $d = 0$ . The solution exists on the open interval  $[0, T_{max})$  with

$$T_{max} = \begin{cases} \infty & \text{if } b \leq -\sqrt{4ac}, \\ \frac{1}{\sqrt{b^2 - 4ac}} \log \frac{b + \sqrt{b^2 - 4ac}}{b - \sqrt{b^2 - 4ac}} & \text{if } b > -\sqrt{4ac}. \end{cases}$$

For  $b \in (-\sqrt{4ac}, \sqrt{4ac}]$ , the function  $b \mapsto \frac{1}{\sqrt{b^2 - 4ac}} \log \frac{b + \sqrt{b^2 - 4ac}}{b - \sqrt{b^2 - 4ac}}$  is to be understood as its analytic continuation out of the domain  $(\sqrt{4ac}, \infty)$ .

*Proof.* The solution formulas are verified by lengthy but straightforward computations. The time horizon  $T_{max}$  is determined by the smallest positive zero of the function  $\tau \mapsto 1 + ge^{-d\tau}$ . ■

**Proof of Theorem 3.4.** Define  $Y_t$  as in (A.6) with (A.7) - (A.9). Since  $z(T-t) > 0$  for all  $t < T$ , there exists a constant  $c > 0$  such that

$$e^{x(T-t) + y(T-t)X_t + z(T-t)X_t^2} \geq e^{x(T-t) - \frac{y(T-t)^2}{4z(T-t)}} \geq c$$

for  $t < T$  and thus  $Y_t \geq cZ_t^2$  for all  $t \in [0, T]$ . Since  $Y_t$  is a continuous process, the stopping times

$$\tau_n = \inf \{ t \in [0, T] \mid Y_t \geq n \} \wedge T$$

satisfy  $\tau_n \nearrow T$  for  $n \rightarrow \infty$ , and the processes  $Y_{t \wedge \tau_n}$  and  $Z_{t \wedge \tau_n}$  are bounded  $P$ -martingales by Lemma A.2 and  $Z_t \leq \sqrt{\frac{1}{c} Y_t}$ . Applying Doob's inequality to  $Z_{t \wedge \tau_n}$ , we find

$$E \left[ \sup_{0 \leq t \leq T} Z_{t \wedge \tau_n}^2 \right] \leq c_2 E [Z_{\tau_n}^2] \leq \frac{c_2}{c} E [Y_{\tau_n}] = \frac{c_2}{c} Y_0$$

for some constant  $c_2 > 0$ . Letting  $n \rightarrow \infty$  and applying monotone convergence in the last inequality, we obtain  $E \left[ \sup_{0 \leq t \leq T} Z_t^2 \right] \leq \frac{c_2}{c} Y_0 < \infty$ . So  $\hat{P} \in \mathcal{M}$ . ■

For the proof of Theorem 3.5 we need the following result.

**Lemma A.4.** Let  $W_t$  be a  $d$ -dimensional Brownian motion on some filtered probability space  $(\Omega, \mathcal{F}, \mathbb{F}, P)$ , and  $a_t$  be an  $\mathbb{R}^d$ -valued and  $b_t, \sigma_t, \nu_t$  be  $\mathbb{R}^{d \times d}$ -valued deterministic functions. Let  $V_t$  be an  $\mathbb{R}^d$ -valued and  $S_t$  an  $\mathbb{R}$ -valued adapted processes satisfying  $S_0 > 0$  and

$$\begin{aligned} dV_t &= (a_t + b_t \cdot V_t)dt + \sigma_t \cdot dW_t, \\ dS_t &= S_t(\nu_t \cdot V_t) \cdot dW_t. \end{aligned}$$

Then  $S_t$  is a martingale.

*Proof.* The proof follows the ideas in Sin [22].  $S_t$  is a positive local martingale and hence a supermartingale, so it suffices to show that  $E[S_T] = S_0$  for each  $T > 0$ . Define the stopping times

$$\tau_n = \inf \{t \geq 0 \mid \int_0^t \|\nu_u \cdot V_u\|^2 du \geq n\}.$$

Since  $\nu_t \cdot V_t$  is a locally bounded process, we have  $\tau_n \nearrow \infty$   $P$ -a.s. for  $n \rightarrow \infty$ . Moreover, the stopped process  $S_t^{\tau_n} = S_{t \wedge \tau_n}$  is a martingale by Novikov's condition. Hence we can define a probability measure  $P^n \approx P$  by  $\frac{dP^n}{dP} = \frac{S_{\tau_n}^{\tau_n}}{S_0^{\tau_n}}$ . Then the process

$$W_t^n = W_t - \int_0^t \nu_u \cdot V_u I_{\{u \leq \tau_n \wedge T\}} du$$

is a  $d$ -dimensional  $P^n$ -Brownian motion by Girsanov's theorem, and  $V_t$  satisfies

$$dV_t = \left( a_t + (\sigma_t \cdot \nu_t I_{\{t \leq \tau_n \wedge T\}} + b_t) \cdot V_t \right) dt + \sigma_t \cdot dW_t^n.$$

Now define a process  $\hat{V}_t$  by  $\hat{V}_0 = V_0$  and

$$d\hat{V}_t = \left( a_t + (\sigma_t \cdot \nu_t I_{\{t \leq T\}} + b_t) \cdot \hat{V}_t \right) dt + \sigma_t \cdot dW_t$$

and a sequence of stopping times  $\hat{\tau}_n$  by

$$\hat{\tau}_n = \inf \{t \geq 0 \mid \int_0^t \|\nu_u \cdot \hat{V}_u\|^2 du \geq n\}.$$

Then the distribution of  $\tau_n$  under  $P^n$  is the same as the distribution of  $\hat{\tau}_n$  under  $P$ . Moreover,  $\hat{\tau}_n \nearrow \infty$   $P$ -a.s. for  $n \rightarrow \infty$  since  $\nu_t \cdot \hat{V}_t$  is locally bounded. Monotone convergence therefore yields

$$\begin{aligned} E[S_T] &= E \left[ \lim_{n \rightarrow \infty} S_T I_{\{\tau_n \geq T\}} \right] \\ &= \lim_{n \rightarrow \infty} E \left[ S_T I_{\{\tau_n \geq T\}} \right] \\ &= S_0 \lim_{n \rightarrow \infty} E \left[ \frac{S_T^{\tau_n}}{S_0^{\tau_n}} I_{\{\tau_n \geq T\}} \right] \\ &= S_0 \lim_{n \rightarrow \infty} E^n \left[ I_{\{\tau_n \geq T\}} \right] \\ &= S_0 \lim_{n \rightarrow \infty} E \left[ I_{\{\hat{\tau}_n \geq T\}} \right] = S_0. \end{aligned}$$

This finishes the proof. ■

Applying Lemma A.4 to  $V_t = X_t$  immediately yields

**Corollary A.5.** Let  $x(\cdot), y(\cdot), z(\cdot)$  be a solution to (A.7) - (A.9) with  $x(0) = y(0) = z(0) = 0$ . Then the process  $\hat{Z}_t = \frac{Y_t}{Z_t}$  in Lemma A.2 is a  $\hat{P}$ -martingale.

**Proof of Theorem 3.5.** a) We proceed in three steps.

*Step 1)* By Lemma 1 c) in [20], it suffices to show that

$$Z_T = M_0 + J_T \tag{A.12}$$

where  $M_0 \in [1, \infty)$  and  $J_T$  is in the  $L^2(P)$ -closure of  $G_T(\Theta)$ , that is in  $G_T(\mathcal{A})$  by Theorem A.1 b). To this end let  $\mathbb{G} = (\mathcal{G}_t)_{t \in [0, T]}$  be the filtration  $\mathcal{G}_t := \sigma((\hat{W}_s^0, \hat{W}_s^1) \mid s \leq t)$  generated by the 2-dimensional  $\hat{P}$ -Brownian motion  $(\hat{W}_t^0, \hat{W}_t^1)$  and define the  $\mathbb{G}$ -stopping times

$$\tau_k = \inf \{t \geq 0 \mid |X_t| \geq k\} \wedge T$$

for  $k \in \mathbb{N}$ . Since  $X_t$  is continuous, we have  $\tau_k \nearrow T$  a.s. for  $k \rightarrow \infty$ . Next define the processes

$$\begin{aligned} M_t &= \frac{1}{Z_t} E[Z_T^2 \mid \mathcal{F}_t] = \hat{E}[Z_T \mid \mathcal{F}_t], \\ M_t^{(k)} &= \frac{1}{Z_t} E[Z_T Z_{\tau_k} \mid \mathcal{F}_t] = \hat{E}[Z_{\tau_k} \mid \mathcal{F}_t]. \end{aligned}$$

Since  $Z_T$  and  $Z_{\tau_k}$  are  $\mathcal{G}_T$ -measurable and  $\mathcal{F}_t = \mathcal{G}_t \vee \sigma(W_s^2 \mid s \leq t)$  with  $W^2 = \hat{W}^2$  independent of  $\mathbb{G}$ , we obtain  $M_t = \hat{E}[Z_T \mid \mathcal{G}_t]$  and  $M_t^{(k)} = \hat{E}[Z_{\tau_k} \mid \mathcal{G}_t]$ , and hence by Itô's representation theorem

$$\begin{aligned} Z_T = M_T = M_0 &+ \int_0^T h_s \cdot d\hat{W}_s, \\ M_t^{(k)} = M_0^{(k)} &+ \int_0^t h_s^{(k)} \cdot d\hat{W}_s, \quad t \in [0, T] \end{aligned}$$

for some predictable processes  $h_t = (h_t^0, h_t^1, 0)$  and  $h_t^{(k)} = (h_t^{(k,0)}, h_t^{(k,1)}, 0)$ . Setting  $J_T = \int_0^T h_s \cdot d\hat{W}_s$ , we obtain (A.12) with  $M_0 = E[Z_T^2] \geq E[Z_T]^2 = 1$ .

*Step 2)* It remains to show that  $J_T$  is in  $G_T(\mathcal{A})$ . To this end recall that  $Z_t$  is a square-integrable  $P$ -martingale by Theorem 3.4, so dominated convergence and Doob's inequality imply that

$$\begin{aligned} M_T^{(k)} = Z_{\tau_k} &\rightarrow Z_T = M_T \quad \text{in } L^2(P), \\ M_0^{(k)} = E[Z_T Z_{\tau_k}] &\rightarrow E[Z_T^2] = M_0 \end{aligned}$$

for  $k \rightarrow \infty$ , and therefore

$$\int_0^T h_s^{(k)} \cdot d\hat{W}_s \rightarrow \int_0^T h_s \cdot d\hat{W}_s = J_T \quad \text{in } L^2(P).$$

Since  $G_T(\mathcal{A})$  is closed in  $L^2(P)$ , it thus suffices to show that  $\int_0^T h_s^{(k)} \cdot d\hat{W}_s \in G_T(\mathcal{A})$  for each  $k$ . To verify this, first note that the nonsingularity of the volatility matrix of  $\vec{F}(t)$  allow us to write

$$\int_0^t h_s^{(k)} \cdot d\hat{W}_s = \int_0^t \zeta_s^{(k)} \cdot d\vec{F}(s) = G_t(\zeta^{(k)}), \quad t \in [0, T]$$

for a suitable predictable and  $\vec{F}$ -integrable process  $\zeta^{(k)}$ . By Theorem A.1 a), the assertion now follows once we show that  $G_t(\zeta^{(k)})$  is a  $Q$ -martingale for each  $Q \in \mathcal{M}$ .

*Step 3)* To this end fix  $k \in \mathbb{N}$  and  $Q \in \mathcal{M}$ . Clearly  $G_t(\zeta^{(k)})$  is a  $Q$ -local martingale. To show the martingale property under  $Q$ , we start by computing

$$Z_{t \wedge \tau_k} = \mathcal{E}\left(-\int \hat{\theta}^{(k)} \cdot dW\right)_t = \mathcal{E}\left(-\int \hat{\theta}^{(k)} \cdot d\hat{W}\right)_t e^{\int_0^t \|\hat{\theta}_s^{(k)}\|^2 ds} = N_t B_t,$$

where  $\hat{\theta}_t^{(k)} = \hat{\theta}_t I_{\{t < \tau_k\}}$  is a process bounded by some constant  $c_k$  depending on  $k$  and the model parameters,  $B_t = e^{\int_0^t \|\hat{\theta}_s^{(k)}\|^2 ds}$ , and  $N_t = \mathcal{E}\left(-\int \hat{\theta}^{(k)} \cdot d\hat{W}\right)_t$  is a  $\hat{P}$ -martingale by Novikov's condition. Hence

$$0 \leq M_0^{(k)} + G_t(\zeta^{(k)}) = M_t^{(k)} = \hat{E}[M_T^{(k)} | \mathcal{F}_t] = \hat{E}[Z_{\tau_k} | \mathcal{F}_t] = \hat{E}[N_T B_T | \mathcal{F}_t] \leq e^{c_k^2 T} N_t \leq e^{c_k^2 T} Z_{t \wedge \tau_k}.$$

It follows that  $\sup_{t \in [0, T]} |G_t(\zeta^{(k)})| \in L^2(P)$  by Theorem 3.4 and Doob's inequality. Hence

$$E^Q \left[ \sup_{t \in [0, T]} |G_t(\zeta^{(k)})| \right] = E \left[ \frac{dQ}{dP} \sup_{t \in [0, T]} |G_t(\zeta^{(k)})| \right] \leq E \left[ \left( \frac{dQ}{dP} \right)^2 \right] E \left[ \left( \sup_{t \in [0, T]} |G_t(\zeta^{(k)})| \right)^2 \right] < \infty$$

by the Cauchy-Schwarz inequality, and so  $G_t(\zeta^{(k)})$  is a  $Q$ -martingale.

b) By a) we have  $\tilde{P} = \hat{P}$  and thus  $\frac{d\tilde{P}}{dP} = Z_T = \hat{Z}_T$ . Since  $\hat{Z}_t$  is a  $\hat{P}$ -martingale by Corollary A.5, it follows that

$$\tilde{Z}_t = \tilde{E} \left[ \frac{d\tilde{P}}{dP} | \mathcal{F}_t \right] = \hat{E}[\hat{Z}_T | \mathcal{F}_t] = \hat{Z}_t.$$

Equation (3.14) now follows from Lemma A.2. For (3.15), note that (2.14) and  $d\hat{W}_t = dW_t + \hat{\theta}_t dt$  imply

$$\begin{aligned} \begin{pmatrix} dF(t, T_i) \\ dF(t, T_h) \end{pmatrix} &= \begin{pmatrix} F(t, T_i)\sigma\rho_0 & F(t, T_i)(\sigma\rho_1 - \eta e_1(\beta, T_i - t)) \\ F(t, T_h)\sigma\rho_0 & F(t, T_h)(\sigma\rho_1 - \eta e_1(\beta, T_h - t)) \end{pmatrix} \begin{pmatrix} d\hat{W}_t^0 \\ d\hat{W}_t^1 \end{pmatrix}, \\ \begin{pmatrix} d\hat{W}_t^0 \\ d\hat{W}_t^1 \end{pmatrix} &= \frac{1}{F(t, T_i)F(t, T_h)w(t)} \begin{pmatrix} F(t, T_h)(\sigma\rho_1 - \eta e_1(\beta, T_h - t)) & F(t, T_i)(-\sigma\rho_1 + \eta e_1(\beta, T_i - t)) \\ -F(t, T_h)\sigma\rho_0 & F(t, T_i)\sigma\rho_0 \end{pmatrix} \\ &\quad \cdot \begin{pmatrix} dF(t, T_i) \\ dF(t, T_h) \end{pmatrix}. \end{aligned} \tag{A.13}$$

Plugging this into (3.14) yields (3.15). ■

**Proof of Theorem 3.6.** We give the proof under the assumption  $\phi > \beta > 0$ , which is satisfied for the parameter estimates we find in our calibration procedure. The result can be easily extended to general parameter values of  $\phi$  and  $\beta$ .

a) By definition of the spot-futures spread in (2.1) with  $T_1 = T_1(t)$ , we have

$$e^{-ru} S_u = e^{-T_1(u)r} F(u, T_1(u)) e^{Y_u}.$$

From (3.16) we then compute

$$\begin{aligned} V_t(\lambda) &= \tilde{E}[H(\lambda) | \mathcal{F}_t] \\ &= \lambda - X_0 + \int_0^t e^{-ru} S_u du + \int_t^T \tilde{E}[e^{-T_1(u)r} F(u, T_1(u)) e^{Y_u} | \mathcal{F}_t] du \\ &= \lambda - X_0 + \int_0^t e^{-ru} S_u du + \sum_{j=1}^k e^{-T_j r} \int_{T_{j-1}}^{T_j} \tilde{E}[F(u, T_j) e^{Y_u} | \mathcal{F}_t] du. \end{aligned} \tag{A.14}$$

Fix  $u \in [0, T]$  and  $T_j$ . We claim that

$$\tilde{E}[F(u, T_j) e^{Y_u} | \mathcal{F}_t] = F(t, T_j) e^{m_j(u-t) + n_1(u-t)Y_t + n_2(u-t)X_t}, \quad t \leq u, \tag{A.15}$$

for suitable deterministic functions  $m_j(\tau), n_1(\tau), n_2(\tau)$  with  $m_j(0) = n_2(0) = 0$  and  $n_1(0) = 1$ . Indeed,

applying Itô's formula to  $M_t^j(u) := F(t, T_j)e^{m_j(u-t)+n_1(u-t)Y_t+n_2(u-t)X_t}$ , and using

$$dF(t, T_j) = F(t, T_j) \left( \sigma \rho_0 d\hat{W}_t^0 + (\sigma \rho_1 - \eta e_1(\beta, T_j - t)) d\hat{W}_t^1 \right), \quad (\text{A.16})$$

$$dX_t = -\kappa X_t dt + \eta dW_t^1 = (-\beta_1 \eta - \beta X_t) dt + \eta d\hat{W}_t^1,$$

$$dY_t = \phi(b - Y_t) dt + \nu(c_0 dW_t^0 + c_1 dW_t^1 + c_2 dW_t^2)$$

$$= \left( \phi b - \nu(\beta_0 c_0 + \beta_1 c_1) - \nu(\gamma_0 c_0 + \gamma_1 c_1) X_t - \phi Y_t \right) dt + \nu(c_0 d\hat{W}_t^0 + c_1 d\hat{W}_t^1 + c_2 d\hat{W}_t^2),$$

from (2.14), (2.8), and (2.15), we find (writing  $n_1 = n_1(u - t)$ ,  $n_1 = n_1(u - t)$  and  $m_j = m_j(u - t)$ )

$$\begin{aligned} dM_t^j(u) = & M_t^j(u) \left( -m_j' - n_1' Y_t - n_2' X_t + n_1 \left( \phi b - \nu(\beta_0 c_0 + \beta_1 c_1) - \nu(\gamma_0 c_0 + \gamma_1 c_1) X_t - \phi Y_t \right) \right. \\ & + n_2 \left( -\beta_1 \eta - \beta X_t \right) + \frac{1}{2} n_1^2 \nu^2 + \frac{1}{2} n_2^2 \eta^2 + n_1 \left( \sigma \rho_0 \nu c_0 + (\sigma \rho_1 - \eta e_1(\beta, T_j - t)) \nu c_1 \right) \\ & \left. + n_2 (\sigma \rho_1 - \eta e_1(\beta, T_j - t)) \eta + n_1 n_2 \eta \nu c_1 \right) dt \\ & + M_t^j(u) \left( (\sigma \rho_0 + c_0 \nu n_1) d\hat{W}_t^0 + \left( (\sigma \rho_1 - \eta e_1(\beta, T_j - t)) + c_1 \nu n_1 + \eta n_2 \right) d\hat{W}_t^1 + c_2 \nu n_1 d\hat{W}_t^2 \right). \end{aligned} \quad (\text{A.17})$$

Hence the drift of  $M_t^j(u)$  is zero if  $m_j, n_1, n_2$  satisfy the ODE system

$$n_1' = -\phi n_1,$$

$$n_2' = -\nu(\gamma_0 c_0 + \gamma_1 c_1) n_1 - \beta n_2,$$

$$\begin{aligned} m_j' = & \left( \phi b - \nu(\beta_0 c_0 + \beta_1 c_1) + \sigma \rho_0 \nu c_0 + (\sigma \rho_1 - \eta e_1(\beta, T_j - t)) \nu c_1 \right) n_1 \\ & + \left( -\beta_1 \eta + (\sigma \rho_1 - \eta e_1(\beta, T_j - t)) \eta \right) n_2 + \frac{1}{2} \nu^2 n_1^2 + \frac{1}{2} \eta^2 n_2^2 + \eta \nu c_1 n_1 n_2 \end{aligned}$$

with  $m_j(0) = n_2(0) = 0$  and  $n_1(0) = 1$ , and lengthy but straightforward calculations show that the solution to this system is given by

$$n_1(\tau) = e^{-\phi\tau}, \quad (\text{A.18})$$

$$n_2(\tau) = \alpha(e^{-\phi\tau} - e^{-\beta\tau}), \quad (\text{A.19})$$

$$m_j(\tau) = \frac{k_1}{\phi}(1 - e^{-\phi\tau}) + \frac{k_2}{2\phi}(1 - e^{-2\phi\tau}) + \frac{k_3}{\beta}(1 - e^{-\beta\tau}) + \frac{k_4}{2\beta}(1 - e^{-2\beta\tau}) + \frac{k_5}{\phi + \beta}(1 - e^{-(\phi + \beta)\tau}), \quad (\text{A.20})$$

where  $\alpha = \frac{\nu(\gamma_0 c_0 + \gamma_1 c_1)}{\phi - \beta}$  and

$$k_1 = \phi b - \nu(\beta_0 c_0 + \beta_1 c_1) + \nu \sigma (\rho_0 c_0 + \rho_1 c_1) - \frac{c_1 \eta \nu}{\beta} + \eta (\sigma \rho_1 - \frac{\eta}{\beta} - \beta_1) \alpha,$$

$$k_2 = \frac{1}{2} \nu^2 + \frac{1}{2} \eta^2 \alpha^2 + c_1 \eta \nu \alpha,$$

$$k_3 = -\eta (\sigma \rho_1 - \frac{\eta}{\beta} - \beta_1) \alpha,$$

$$k_4 = \frac{1}{2} \eta^2 \alpha^2 - \frac{\eta^2}{\beta} e^{-\beta(T_j - u)} \alpha,$$

$$k_5 = -\eta^2 \alpha^2 + \left( \frac{c_1 \nu \eta}{\beta} + \frac{\eta^2}{\beta} \alpha \right) e^{-\beta(T_j - u)} - c_1 \nu \eta \alpha.$$

In this case,  $M_t^j(u)$  is a  $\tilde{P}$ -local martingale, and since the diffusion coefficient is of the form  $M_t^j(u)c(t)$  with a (deterministic) bounded function  $c(t)$ , the process  $M_t^j(u)$  is a  $\tilde{P}$ -martingale by Novikov's condition. Now (A.15) follows from  $M_u^j(u) = F(u, T_j)e^{Y_u}$ . Together with (A.14) we obtain (3.17).

b) Plugging  $M_t^j(u) = F(t, T_j)q_j(t, u)$  into (A.17) and then using (A.16), we obtain

$$dM_t^j(u) = q_j(t, u) dF(t, T_j) + F(t, T_j)q_j(t, u) \left( c_0 \nu n_1 d\hat{W}_t^0 + (c_1 \nu n_1 + \eta n_2) d\hat{W}_t^1 + c_2 \nu n_1 d\hat{W}_t^2 \right). \quad (\text{A.21})$$

Moreover by (3.17) we have

$$V_t(\lambda) = \lambda - X_0 + \int_0^t e^{-ru} S_u du + \sum_{j=1}^k e^{-T_j r} \int_{T_{j-1}}^{T_j} M_t^j(u) du.$$

Applying Itô's formula here and using that  $V_t(\lambda)$  and  $M_t^j(u)$  are  $\tilde{P}$ -martingales, it follows that

$$dV_t(\lambda) = \sum_{j=1}^k e^{-T_j r} \int_{T_{j-1}}^{T_j} dM_t^j(u) du.$$

Plugging in (A.21) here, we obtain

$$\begin{aligned} dV_t(\lambda) &= \sum_{j=1}^k e^{-T_j r} \int_{T_{j-1}}^{T_j} q_j(t, u) du dF(t, T_j) \\ &\quad + \left( \sum_{j=1}^k e^{-T_j r} F(t, T_j) \int_{T_{j-1}}^{T_j} q_j(t, u) c_0 \nu n_1 (u - t) du \right) d\hat{W}_t^0 \\ &\quad + \left( \sum_{j=1}^k e^{-T_j r} F(t, T_j) \int_{T_{j-1}}^{T_j} q_j(t, u) (c_1 \nu n_1 (u - t) + \eta n_2 (u - t)) du \right) d\hat{W}_t^1 \\ &\quad + dL_t \\ &= \sum_{j=1}^k e^{-T_j r} \int_{T_{j-1}}^{T_j} q_j(t, u) du dF(t, T_j) + C_t^0 d\hat{W}_t^0 + C_t^1 d\hat{W}_t^1 + dL_t \end{aligned}$$

with a  $\tilde{P}$ -local martingale  $L_t$  orthogonal to  $\vec{F}(t)$ . Plugging (A.13) into the last equation yields (3.18). ■

**Proof of Theorem 3.7.** The structure of the proof is analogous to the proof of Theorem 3.6 a), so we only give a sketch. As in (A.14) we obtain that

$$E[H_T] = - \sum_{j=1}^k e^{-T_j r} \int_{T_{j-1}}^{T_j} E[F(u, T_j) e^{Y_u}] du,$$

so (3.21) follows once we show that for all  $t \leq u$

$$E[F(u, T_j) e^{Y_u} | \mathcal{F}_t] = F(t, T_j) e^{\ell_j(u-t) + p(u-t)Y_t + s_j(u-t)X_t} \quad (\text{A.22})$$

with deterministic functions  $p, \ell_j, s_j$  satisfying  $p(0) = 1$  and  $s_j(0) = \ell_j(0) = 0$ . To this end, we apply Itô's formula to the RHS of (A.22), use (2.14), (2.8), and (2.15), and as in the proof of Theorem 3.6 a), we find that the RHS of (A.22) is a  $P$ -local martingale, and then indeed a martingale, if and only if the functions  $p, s_j, \ell_j$  fulfill a system of ODEs. This system can be solved explicitly, and lengthy but

straightforward computations yield that

$$p(\tau) = e^{-\phi\tau}, \quad (\text{A.23})$$

$$s_j(\tau) = -\alpha - \frac{1}{\beta}e^{-\beta(T_j-u)}e^{-\beta\tau} + \left(\alpha + \frac{1}{\beta}e^{-\beta(T_j-u)}\right)e^{-\kappa\tau}, \quad (\text{A.24})$$

$$\begin{aligned} \ell_j(\tau) = & k_0\tau + \frac{k_1}{\beta}(1 - e^{-\beta\tau}) + \frac{k_2}{\phi}(1 - e^{-\phi\tau}) + \frac{k_3}{\kappa}(1 - e^{-\kappa\tau}) \\ & + \frac{k_4}{2\beta}(1 - e^{-2\beta\tau}) + \frac{k_5}{2\phi}(1 - e^{-2\phi\tau}) + \frac{k_6}{2\kappa}(1 - e^{-2\kappa\tau}) + \frac{k_7}{\phi + \kappa}(1 - e^{-(\phi+\kappa)\tau}), \end{aligned} \quad (\text{A.25})$$

where  $\alpha = \frac{\eta\gamma_1}{\beta\kappa} - \frac{\sigma}{\kappa}(\rho_0\gamma_0 + \rho_1\gamma_1)$  and

$$\begin{aligned} k_0 &= (\sigma\rho_0\beta_0 + \sigma\rho_1\beta_1 - \frac{\eta\beta_1}{\beta}) - \eta(\sigma\rho_1 - \frac{\eta}{\beta})\alpha + \frac{1}{2}\eta^2\alpha^2, \\ k_1 &= \frac{\eta}{\beta}(\beta_1 - \sigma\rho_1 + \frac{\eta}{\beta})e^{-\beta(T_j-u)}, \\ k_2 &= \phi b + \sigma\nu(c_0\rho_0 + c_1\rho_1) - \frac{c_1\nu\eta}{\beta} - c_1\eta\nu\alpha, \\ k_3 &= \left(\eta(\sigma\rho_1 - \frac{\eta}{\beta}) - \eta^2\alpha\right)\left(\alpha + \frac{1}{\beta}e^{-\beta(T_j-u)}\right), \\ k_4 &= -\frac{1}{2}\frac{\eta^2}{\beta^2}e^{-2\beta(T_j-u)}, \\ k_5 &= \frac{1}{2}\nu^2, \\ k_6 &= \frac{1}{2}\eta^2\left(\alpha + \frac{1}{\beta}e^{-\beta(T_j-u)}\right)^2, \\ k_7 &= c_1\eta\nu\left(\alpha + \frac{1}{\beta}e^{-\beta(T_j-u)}\right). \end{aligned}$$

To verify (3.22), similarly as above we compute

$$E[H_T^2] = \sum_{i=1}^k \sum_{j=1}^k e^{-(T_i+T_j)r} \int_{T_{i-1}}^{T_i} \left( \int_{T_{j-1}}^{T_j} E[F(u, T_i)F(v, T_j)e^{Y_u+Y_v}] dv \right) du.$$

Hence the assertion follows once we show

$$E[F(u, T_i)F(v, T_j)e^{Y_u+Y_v}] = F(0, T_i)F(0, T_j)q_{ij}(u, v) \quad (\text{A.26})$$

for all  $u, v$ , and it suffices to establish (A.26) for  $u \geq v$  by symmetry of the function  $q_{ij}(u, v)$  in  $u$  and  $v$ . So let  $u \geq v$ . We note that by (A.22) we have

$$\begin{aligned} E[F(u, T_i)F(v, T_j)e^{Y_u+Y_v}] &= E\left[F(v, T_j)e^{Y_v} E[F(u, T_i)e^{Y_u} \mid \mathcal{F}_v]\right] \\ &= E\left[F(v, T_j)e^{Y_v} F(v, T_i)e^{\ell_i(u-v)+p(u-v)Y_v+s_i(u-v)X_v}\right] \\ &= E\left[F(v, T_j)F(v, T_i)e^{(1+p(u-v))Y_v+s_i(u-v)X_v}\right]e^{\ell_i(u-v)}, \end{aligned}$$

and thus (A.26) follows once we prove for all  $t \in [0, v]$

$$E\left[F(v, T_j)F(v, T_i)e^{(1+p(u-v))Y_v+s_i(u-v)X_v} \mid \mathcal{F}_t\right] = F(t, T_j)F(t, T_i)e^{m_{ij}(v-t)+w(v-t)Y_t+w_{ij}(v-t)X_t} \quad (\text{A.27})$$

with deterministic functions  $w, w_{ij}, m_{ij}$  satisfying the equations  $m_{ij}(0) = 0$ ,  $w(0) = 1 + p(u - v)$ , and  $w_{ij}(0) = s_i(u - v)$ . To this end, we proceed as above. We apply Itô's formula to the RHS of (A.27), use (2.14), (2.8), and (2.15), and as in the proof of Theorem 3.6 a), we find that the RHS of (A.27) is a



$P$ -local martingale, and then indeed a martingale, if and only if the functions  $w, w_{ij}, m_{ij}$  fulfill a system of ODEs. This system can be solved explicitly, and lengthy but straightforward computations yield that

$$w(\tau) = (1 + p(u - v))e^{-\phi\tau}, \quad (\text{A.28})$$

$$w_{ij}(\tau) = -2\alpha - \frac{1}{\beta}(e^{-\beta(T_i-v)} + e^{-\beta(T_j-v)})e^{-\beta\tau} + \left(2\alpha + \frac{1}{\beta}(e^{-\beta(T_i-v)} + e^{-\beta(T_j-v)}) + s_i(u - v)\right)e^{-\kappa\tau}, \quad (\text{A.29})$$

$$m_{ij}(\tau) = k_0\tau + \frac{k_1}{\beta}(1 - e^{-\beta\tau}) + \frac{k_2}{\phi}(1 - e^{-\phi\tau}) + \frac{k_3}{\kappa}(1 - e^{-\kappa\tau}) + \frac{k_4}{2\beta}(1 - e^{-2\beta\tau}) + \frac{k_5}{2\phi}(1 - e^{-2\phi\tau}) + \frac{k_6}{2\kappa}(1 - e^{-2\kappa\tau}) + \frac{k_7}{\phi + \kappa}(1 - e^{-(\phi+\kappa)\tau}), \quad (\text{A.30})$$

where  $p(\cdot)$  and  $s_i(\cdot)$  are defined in (A.23), (A.24),  $\alpha = \frac{\eta\gamma_1}{\beta\kappa} - \frac{\sigma}{\kappa}(\rho_0\gamma_0 + \rho_1\gamma_1)$ , and

$$\begin{aligned} k_0 &= 2(\sigma\rho_0\beta_0 + \sigma\rho_1\beta_1 - \frac{\eta\beta_1}{\beta}) + \sigma^2\rho_0^2 - 4\eta(\sigma\rho_1 - \frac{\eta}{\beta})\alpha + 2\eta^2\alpha^2, \\ k_1 &= \frac{\eta}{\beta}(\beta_1 - \sigma\rho_1 + \frac{\eta}{\beta})(e^{-\beta(T_i-v)} + e^{-\beta(T_j-v)}), \\ k_2 &= \left(\phi b + 2\sigma\nu(c_0\rho_0 + c_1\rho_1) - 2\frac{c_1\nu\eta}{\beta} - 2c_1\eta\nu\alpha\right)(1 + p(u - v)), \\ k_3 &= 2\left(\eta(\sigma\rho_1 - \frac{\eta}{\beta}) - \eta^2\alpha\right)\left(2\alpha + \frac{1}{\beta}(e^{-\beta(T_i-v)} + e^{-\beta(T_j-v)}) + s_i(u - v)\right), \\ k_4 &= \frac{\eta^2}{\beta^2}e^{-\beta(T_i+T_j-2v)} - \frac{1}{2}\frac{\eta^2}{\beta^2}(e^{-\beta(T_i-v)} + e^{-\beta(T_j-v)})^2, \\ k_5 &= \frac{1}{2}\nu^2(1 + p(u - v))^2, \\ k_6 &= \frac{1}{2}\eta^2\left(2\alpha + \frac{1}{\beta}(e^{-\beta(T_i-v)} + e^{-\beta(T_j-v)}) + s_i(u - v)\right)^2, \\ k_7 &= c_1\eta\nu(1 + p(u - v))\left(2\alpha + \frac{1}{\beta}(e^{-\beta(T_i-v)} + e^{-\beta(T_j-v)}) + s_i(u - v)\right). \end{aligned}$$

This finishes the proof. ■

## References

- [1] M. Bertus, J. Godbey, J. Hilliard, *Minimum variance cross hedging under mean-reverting spreads, stochastic convenience yields, and jumps: Application to the airline industry*, Journal of Futures Markets 29, 736-756 (2009)
- [2] R. Caldentey and M. Haugh, *Optimal Control and Hedging of Operations in the Presence of Financial Markets*, Mathematics of Operations Research 31, 285-304 (2006)
- [3] R. Carmona and M. Ludkovski, *Spot Convenience Yield Models for the Energy Markets*, in: G. Yin and Q. Zhang (eds.), *Mathematics of Finance*, AMS Comm. 351, 65-80 (2004)
- [4] S.-S. Chen, C.-f. Lee, K. Shrestha, *Futures hedge ratios: a review*, The Quarterly Review of Economics and Finance 43, 433-465 (2003)
- [5] A. Černý and J. Kallsen, *On the structure of general mean-variance hedging strategies*, Annals of Probability 35, 1479-1531 (2007)
- [6] A. Černý and J. Kallsen, *Mean-variance hedging and optimal investment in Heston's model with correlation*, Mathematical Finance 18, 473-492 (2008)

- [7] F. Delbaen and W. Schachermayer, *The variance-optimal martingale measure for continuous processes*, Bernoulli 2, 81-105 (1996)
- [8] F. Delbaen and W. Schachermayer, *Attainable claims with  $p$ -th moments*, Annales de l'Institut Henri Poincaré 32, 743-763 (1996)
- [9] D. Duffie, J. Pan, and K. Singleton, *Transform analysis and asset pricing for affine jump-diffusions*, Econometrica. Journal of the Econometric Society 68, 1343-1376 (2000)
- [10] H. Föllmer and M. Schweizer, *Hedging of contingent claims under incomplete information*, in: M. H. A. Davis and R. J. Elliott (eds.), *Applied Stochastic Analysis*, Stochastics Monographs, Vol. 5, Gordon and Breach, London, 389-414 (1991)
- [11] C. Gourieroux, J. Laurent, and H. Pham, *Mean-variance hedging and numéraire*, Math. Finance 8, 179-200 (1998)
- [12] D. Lien, Y. Tse, *Some recent developments in futures hedging*, Journal of Economic Surveys 16, 357-383 (2002)
- [13] K. Miltersen and E. Schwartz, *Pricing of options on commodity futures with stochastic term structures of convenience yields and interest rates*, Journal of Financial and Quantitative Analysis 33, 33-59 (1998)
- [14] J. Nascimento and W. Powell, *An Optimal Solution to a General Dynamic Jet Fuel Hedging Problem* (2008), available at <http://www.castlelab.princeton.edu/Papers/NascimentoPowell-JetFuelHedging.pdf>
- [15] Pham, H., *Continuous-time stochastic control and optimization with financial applications*, Springer, Berlin (2009)
- [16] H. Pham, T. Rheinländer and M. Schweizer, *Mean-Variance Hedging for Continuous Processes: New Results and Examples*, Finance and Stochastics 2, 173-198 (1998)
- [17] E. Platen, *On the role of the growth optimal portfolio in finance*, Australian Economic Papers 44, 365-388 (2005)
- [18] T. Rheinländer and M. Schweizer, *On  $L^2$ -projections on a space of stochastic integrals*, Annals of Probability 25, 1810-1831 (1997)
- [19] M. Schweizer, *On the Minimal Martingale Measure and the Föllmer-Schweizer Decomposition*, Stochastic Analysis and Applications 13, 573-599 (1995)
- [20] M. Schweizer, *Approximation Pricing and the Variance-Optimal Martingale Measure*, Annals of Probability 24, 206-236 (1996)
- [21] M. Schweizer, *A guided tour through quadratic hedging approaches*, in: E. Jouini, J. Cvitanic, M. Musiela (eds.), *Option Pricing, Interest Rates and Risk Management*, Cambridge University Press, 538-574 (2001)
- [22] C. Sin, *Complications with stochastic volatility models*, Advances in Applied Probability 30, 256-268 (1998)