

Futures Cross-hedging with a Stationary Basis

Stefan Ankirchner Georgi Dimitroff Gregor Heyne
Christian Pigorsch*

When managing risk, frequently only imperfect hedging instruments are at hand. We show how to optimally *cross-hedge* risk when the spread between the hedging instrument and the risk is *stationary*. At the short end, the optimal hedge ratio is close to the cross-correlation of the log returns, whereas at the long end, it is optimal to fully hedge the position. For linear risk positions we derive explicit formulas for the hedge error, and for non-linear positions we show how to obtain numerically efficient estimates. Finally, we demonstrate that even in cases with no clear-cut decision concerning the stationarity of the spread it is better to allow for *mean reversion of the spread* rather than to neglect it.

This version : August 30, 2011

JEL Classification : C30, C51, G13

Keywords : risk management, cross-hedging, cointegration, futures contracts, continuous-time model

*Ankirchner, Institute for Applied Mathematics, University of Bonn, Endenicher Allee 60, 53113 Bonn, Germany. *Email*: ankirchner@iam.uni-bonn.de, <http://www.iam.uni-bonn.de/people/ankirchner/>, *Phone*: +49 (0)228 73 3796; Dimitroff, DZ Bank AG, Platz der Republik, 60265 Frankfurt am Main, Germany, *Email*: dimitroff@gmail.com; Heyne, Department of Mathematics, Humboldt-Universität zu Berlin, Unter den Linden 6, 10099 Berlin, Germany, *Email*: heyne@math.hu-berlin.de, <http://www.mathematik.hu-berlin.de/~heyne/>, *Phone*: +49 (0)30 2093 5495; Pigorsch, Department of Economics, University of Bonn, Adenauerallee 24-42, 53113 Bonn, Germany. *Email*: christian.pigorsch@uni-bonn.de, <http://ect-pigorsch.ect.uni-bonn.de/>, *Phone*: +49 (0)228 73 3920. We are grateful to Hendrik Bessembinder (the editor), Michael Monoyios, an anonymous referee and seminar participants at Humboldt University Berlin, Technical University Munich and University of Oxford for helpful comments. S.A. and C.P. gratefully acknowledge financial support by the German Research Foundation (DFG) through the *Hausdorff Center for Mathematics*.

I. Introduction

Reducing or even eliminating a particular risk is an important task in risk management. However, often only imperfect hedging instruments are at hand, leading to *basis risk*. This is for instance the case if the asset that is hedged does not exactly coincide with the asset underlying the futures contract. A typical example for such a case is an airline company that wants to protect itself against changing kerosene prices. Since there is no liquid kerosene futures market the airline company may fall back on futures on less refined oil, such as crude oil futures, for hedging its kerosene risk. This is a reasonable approach, if the price evolvments of kerosene and of crude oil are very similar. The upper left panel of Figure 1 illustrates the close comovement of the two price series at the IntercontinentalExchange (ICE).

The correlation between the price changes is the crucial determinant of an optimal cross-hedge. A common approach in the literature and in practice is to obtain the optimal hedge ratio by using the most frequent returns or price increments being available, irrespective of the time to maturity. This is a valid approach if the correlation (between the returns or price increments) and the ratio of the standard deviations are constant with respect to the sampling frequency, such as for correlated (geometric) Brownian motions. However, in many cases the correlation depends strongly on the selected time interval. For example in our empirical illustration the correlation of the daily log returns of kerosene and crude oil is only 0.52, which seems unexpectedly low given the strong comovement in the price series. The correlations of the weekly, monthly and yearly log returns in contrast are at 0.72, 0.84 and 0.98, respectively. Thus, the short-term correlation is considerably lower than the long-term correlation, pointing towards a long-term relationship with potential short-term deviations. This property is closely related to the concept of cointegration. It dates back to Engle and Granger (1987) and Granger (1981) and assumes that a set of time series share a long-term relationship with temporary deviations from this “equilibrium”. More precisely, consider two integrated time series (of order

one). They are cointegrated if a linear combination of them is stationary. This is supported for our example in Figure 1, which shows on the lower panel a clear mean reverting behavior of the spread between the logarithmic prices of kerosene and crude oil. Note that we do not use an estimated cointegration vector but rather assume that the spread between the log prices is stationary. This is more restrictive, but empirically supported by the p -value of the augmented Dickey-Fuller test (which is ≤ 0.001) indicating that the null hypothesis of a non-stationary spread is rejected.

From a fundamental point of view the spread of the kerosene and crude oil price is determined by the marginal costs of producing (the finer) kerosene out of crude oil. Temporary deviations of the spread from the marginal costs may occur due to a kerosene shortage or an oversupply. The speed with which the spread reverts to a mean level essentially reflects how fast the market can compensate the deviations. The common stochastic trend reflects shocks, such as, for example, a natural or political crisis in the producing countries, that affect the price of crude oil and, thus, also affect indirectly the price of kerosene due to the intensive and essential use of crude oil in the production process for kerosene. Kerosene and crude oil, however, is only one example for a pair of cointegrated processes and there are many studies pointing towards a cointegration relation between asset prices and corresponding hedging instruments, see e.g. Alexander (1999), Baillie (1989), Brenner and Kroner (1995), Lien and Luo (1993) and Ng and Pirrong (1996) and the references therein.

The long-term relationship between the kerosene price and the crude oil price leads to the observed increasing correlation in our example so that the optimal hedge ratios are not constant, but depend on time to maturity. Intuitively, for long-term hedges it is likely that the two assets are in their equilibrium relationship, whereas in the short-term the dynamics are dominated by noisy fluctuations due to shortage or oversupply of kerosene. To account for a time varying hedge ratio a possible strategy is to estimate the optimal hedge ratio for different maturity times. However, this strategy is not consistent as the objective function, e.g. the variance of

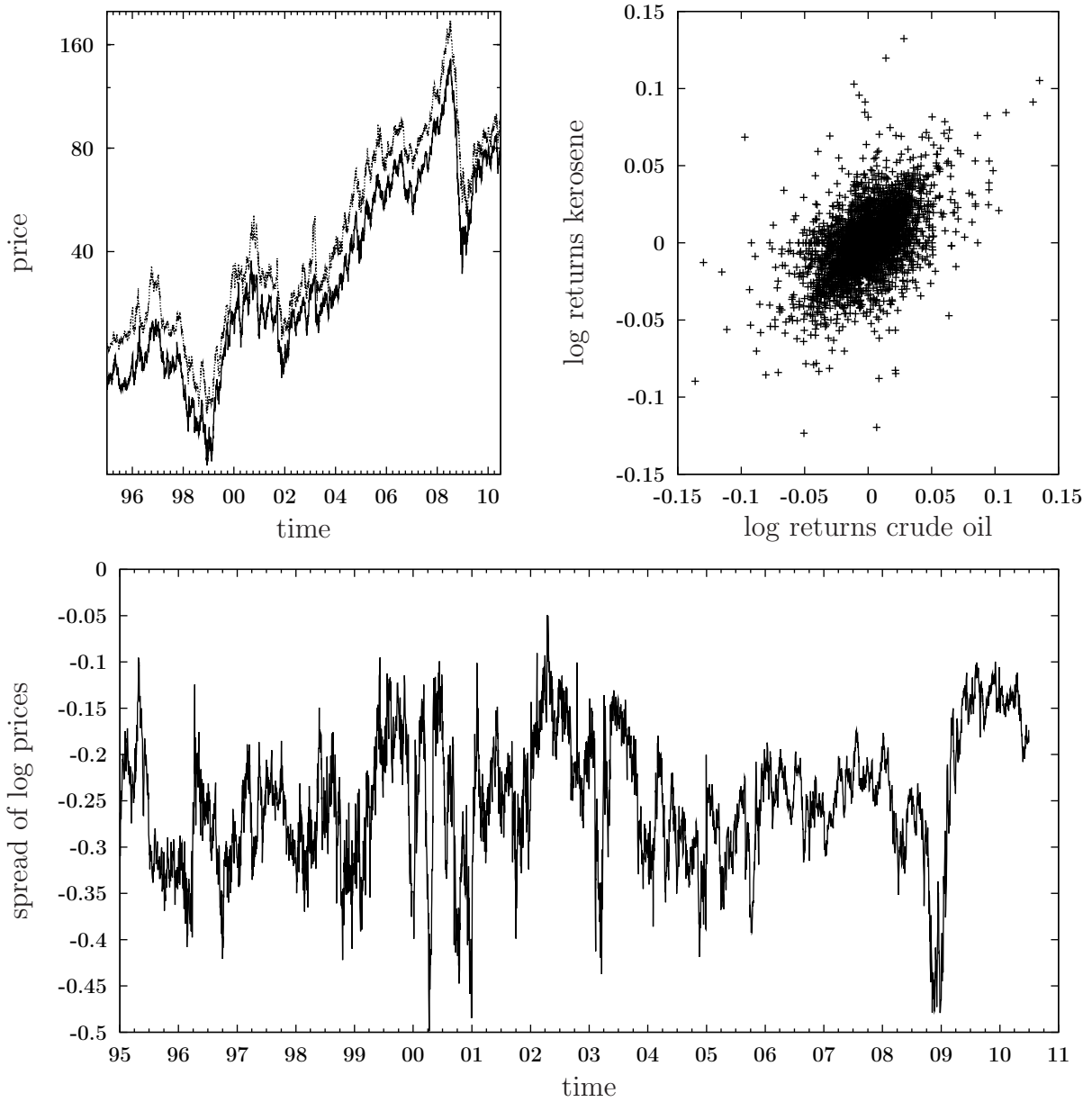


Figure 1: The upper left panel depicts the time evolution of the daily price of crude oil in US\$/BBL (dashed line) and for jet kerosene in US\$/BBL (solid line) from 1995/01/02 until 2010/06/30 (resulting in 4043 observations). The upper right panel exhibits the scatter plot of the corresponding daily log returns and shows that there is positive correlation among the two series as already mentioned in the text (with a correlation coefficient of 0.52). The lower panel depicts the time evolution of the spread of the log prices. Note that crude oil is usually traded in units of barrel, a volume unit. In contrast the trading unit of kerosene is the metric ton, i.e. a weight unit. Although a common unit is not necessary for our analysis we convert the price of kerosene per metric ton to a price per barrel by assuming a density of kerosene of 810 kg/m^3 (at 20 centigrade and 1 atm). This means a metric ton of kerosene has a volume of $(1000/810) \text{ m}^3$ or 1234.6 liter. A barrel of crude oil contains 158.987 liter so that the ratio is 0.1288.

the hedge error, is optimized within a one-period model (with different times to maturity). But the reestimation implies that this strategy is a multiperiod strategy and it therefore results in a suboptimal rule. This is also supported by Lien (1992), who shows that even in a very simple model the multiperiod optimal hedge ratio differs from the hedge ratio obtained in a one-period model.

Therefore, applying a one-period model is imperfect as it does not account for a dynamical rebalance of the hedge. Howard and D'Antonio (1991), Lien (1992, 2004) and Lien and Luo (1993) were among the first to consider a discrete-time multiperiod planning horizon of the agent with (at least partially) cointegrated time series.¹

A natural and self-evident generalization is to consider the possibility of adjustments to the hedging position in continuous-time. This has the additional advantage that it is often possible to compute the standard deviation of the hedge error for different hedging strategies in *closed-form*, which seems to be impossible for discrete-time models. The availability of the standard deviation of the hedge error in turn allows to assess the value of a hedging strategy. Since portfolio risks are usually assessed daily, following a mark-to-market procedure, efficient risk calculation algorithms are needed that allow to estimate the risk quickly. Fast hedge error estimation algorithms are also needed for pricing. If a derivative can only be cross-hedged, then a writer will ask for a premium for the hedge error. In order to determine the risk premia, traders need to quickly estimate expected hedge errors. Obviously, the most convenient support is given by closed-form formulas.

¹Another strand of the literature is based on the recursive estimation of the parameters or states of a “dynamic” model. Among others, Baillie and Myers (1991), Brenner and Kroner (1995) and Cecchetti, Cumby, and Figlewski (1988) consider bivariate (G)ARCH models to estimate the time varying variances and covariances. Although these models account for time variations in the distribution the variance of the hedge is still minimized within a one-period model and with a fixed time to maturity.

The aim of this paper is to set up a model that allows a rigorous study of the effect of a long-term relationship on optimal cross-hedging strategies, and at the same time allows an efficient calculation of the basis risk entailed by the optimal cross-hedges. We choose a continuous-time line, and reproduce the long-term relationship of the prices by describing the spread as an Ornstein-Uhlenbeck process, a Gaussian mean-reverting process, and by modeling the futures price as a geometric Brownian motion (GBM). Noteworthy, our model differs from the widely studied models where both processes, the risk source and the hedging instrument, are GBMs.² In these models, in the following referred to as 2GBM models, the spread of the log prices is not stationary since the variance of the spread is proportional to time. We further show that these models underhedge the risk when cointegration is present (see Section A). Our model, in contrast, explicitly accounts for a stationary spread. Furthermore, it is easy to estimate and it is still tractable enough to allow for a quick calculation of the hedge error standard deviation under different trading strategies. In particular, we are able to derive time-consistent strategies, allowing for updating, that minimize the variance of the hedge error.

To this end, we first solve the optimization problem of finding the dynamic strategy that minimizes the variance of the hedge error. In an abstract continuous martingale setting of an incomplete market, variance optimal hedging strategies have been first described in Föllmer and Sondermann (1986). We make use of their method and transfer it to the specific case of cross-hedging risk with futures contracts within our Markovian model. The optimal hedging strategy can be expressed in terms of the risk's *Greeks* and a hedge ratio decaying with time to maturity. Moreover, for linear risk positions we are able to derive a closed-form formula for

²Such models are considered for example in Ankirchner and Heyne (2011), Duffie and Richardson (1991), Schweizer (1992), who derive cross-hedging strategies minimizing quadratic objective functions, and in Ankirchner, Imkeller, and Dos Reis (2010), Davis (2006), Monoyios (2004), Musiela and Zariphopoulou (2004), who provide cross-hedging strategies maximizing the hedger's expected utility.

the hedge error standard deviation.

The paper is structured as follows. Section II introduces our model and presents some empirical evidence, while Section III briefly reviews hedging with futures contracts and derives the variance optimal hedging strategy for our model. Section IV develops the implied hedge errors within our model for linear and non-linear positions and Section V compares the hedge errors between different models and (suboptimal) hedging strategies emphasizing the importance of allowing for a stationary spread. An extension of our model to account for stochastic volatility is given in Section VI. Section VII concludes while Appendix A contains some empirical results and Appendix B provides the proofs.

II. The Continuous-time Model with a Stationary Spread

As always in modeling real-world phenomena there is a trade-off between the accuracy of a model and its tractability. We therefore illustrate the implications of a stationary spread between the futures price and the price of an illiquid asset by assuming a simplified and tractable model, which is presented in Subsection A. This approach allows us to derive not only optimal hedging strategies and their implied hedge errors (see Section III), but also to obtain the transition density in closed form, so that the model can straightforwardly be estimated via the efficient maximum likelihood method. An empirical illustration is provided in Subsection B.

A. Model Specification

Let $I = (I_t)_{t \geq 0}$ denote the price process of an illiquid asset, and suppose that an economic agent aims at hedging a position $h(I_T)$, where $h : \mathbb{R} \rightarrow \mathbb{R}$ is a measurable payoff function and $T > 0$ is a fixed time horizon. Furthermore, we assume that there exists a liquidly traded futures contract with price process $X = (X_t)_{t \geq 0}$, which evolves according to the stochastic differential

equation (SDE)

$$(1) \quad dX_t = \mu X_t dt + \sigma_X X_t dW_t^{(X)}, \quad X_0 = x,$$

with volatility $\sigma_X > 0$ and constant drift rate $\mu \in \mathbb{R}$. The process $W^{(X)} = (W_t^{(X)})_{t \geq 0}$ is a Brownian motion on a stochastic basis with probability measure P . We denote the spread of the log prices, in the following simply referred to as the *logspread*, by

$$S_t = \log(X_t) - \log(I_t).$$

Although the non-stationarity of the logspread seems to be a plausible assumption for certain asset classes, e.g. for stock prices, there also exist relevant examples for stationary logspreads as shown in the introduction and the mentioned articles. We therefore propose to account for cointegration by first modeling the logspread as a stationary process and then derive the implied dynamics of the illiquid asset.

More precisely, we assume that the logspread follows a (Gaussian) Ornstein-Uhlenbeck process, which is the continuous-time analogue of the stationary discrete-time first order autoregressive process. Under this assumption the logspread solves the mean reverting SDE

$$(2) \quad dS_t = \kappa(m - S_t) dt + \sigma_S(\rho dW_t^{(X)} + \bar{\rho} dW_t^\perp), \quad S_0 = s,$$

where $W^\perp = (W_t^\perp)_{t \geq 0}$ is a Brownian motion independent of $W^{(X)}$, $\kappa \geq 0$ is the mean reversion speed, and $\rho \in [-1, 1]$ the correlation. The logspread's volatility σ_S is assumed to be non-negative. Moreover, we define $\bar{\rho} = \sqrt{1 - \rho^2}$ and use, for the ease of exposition, the following short-hand notation

$$W_t^{(S)} = \rho W_t^{(X)} + \bar{\rho} W_t^\perp$$

for the Brownian motion driving the logspread. Note that for $\kappa \downarrow 0$ the Ornstein-Uhlenbeck process becomes more and more persistent and in the limit a (scaled and shifted) Brownian motion that is correlated with the Brownian motion of the futures price process.

The dynamics of X and S determine the dynamics of the illiquid asset price, as $I_t = X_t e^{-S_t}$, $t \geq 0$. A straightforward calculation, appealing to Itô's formula, shows that the dynamics of I satisfy

$$dI_t = I_t \left(\frac{1}{2} \sigma_S^2 - \kappa(m - S_t) + \mu - \rho \sigma_S \sigma_X \right) dt + I_t \sigma_I dW_t^{(I)},$$

where $\sigma_I = \sqrt{\sigma_X^2 - 2\rho\sigma_S\sigma_X + \sigma_S^2}$ and $W^{(I)} = \left(W_t^{(I)} \right)_{t \geq 0}$ is a Brownian motion defined by $W_t^{(I)} = \left((\sigma_X - \rho\sigma_S)W_t^{(X)} - \bar{\rho}\sigma_S W_t^\perp \right) / \sigma_I$, $t \geq 0$.

Note that the correlation ρ_{IX} between the Brownian motions driving I and X is given by

$$(3) \quad \rho_{IX} = \frac{1}{\sigma_I} (\sigma_X - \rho\sigma_S),$$

which is non-negative if and only if $\sigma_X \geq \rho\sigma_S$. For fixed parameters ρ and σ_X , we can write ρ_{IX} as a function of σ_S :

$$(4) \quad \rho_{IX}(\sigma_S) = \frac{\sigma_X - \rho\sigma_S}{\sqrt{\sigma_X^2 - 2\rho\sigma_S\sigma_X + \sigma_S^2}}.$$

It can be shown that $\rho_{IX}(\sigma_S)$ is strictly decreasing in σ_S , and hence invertible on \mathbb{R}_+ . For a proof of the following result see Appendix B.

Lemma II.1. *Let $\sigma_X > 0$ and $\rho \in (-1, 1)$. Then the mapping $\mathbb{R}_+ \ni \sigma_S \mapsto \rho_{IX}(\sigma_S)$ is strictly decreasing. Moreover, the logspread's volatility σ_S satisfies*

$$(5) \quad \sigma_S = \sigma_X \frac{\sqrt{1 - \rho_{IX}^2}}{\rho\sqrt{1 - \rho_{IX}^2} + \rho_{IX}\sqrt{1 - \rho^2}}.$$

Observe that $(\log X_t, S_t, \log I_t)$ is a 3-dimensional Gaussian process. Furthermore, it possesses a closed-form transition density and hence efficient maximum likelihood estimation becomes feasible. Indeed, a straightforward calculation shows that the triple $(\log X_t, S_t, \log I_t)$ satisfies

$$\begin{pmatrix} \log X_t \\ S_t \\ \log I_t \end{pmatrix} \bigg| \begin{pmatrix} X_0 \\ S_0 \\ I_0 \end{pmatrix} = \begin{pmatrix} x \\ s \\ xe^{-s} \end{pmatrix} \sim \mathcal{N} \left(\begin{pmatrix} \log x + \left(\mu - \frac{\sigma_X^2}{2} \right) t \\ se^{-\kappa t} + m(1 - e^{-\kappa t}) \\ \log x + \left(\mu - \frac{\sigma_X^2}{2} \right) t - se^{-\kappa t} - m(1 - e^{-\kappa t}) \end{pmatrix}, \mathbf{\Sigma} \right),$$

where

$$\mathbf{\Sigma} = \begin{pmatrix} t\sigma_X^2 & \frac{\rho\sigma_X\sigma_S}{\kappa}(1 - e^{-\kappa t}) & t\sigma_X^2 - \rho\frac{\sigma_X\sigma_X}{\kappa}(1 - e^{-\kappa t}) \\ \frac{\rho\sigma_X\sigma_S}{\kappa}(1 - e^{-\kappa t}) & \frac{\sigma_S^2}{2\kappa}(1 - e^{-2\kappa t}) & \frac{\rho\sigma_X\sigma_S}{\kappa}(1 - e^{-\kappa t}) - \frac{\sigma^2}{2\kappa}(1 - e^{-2\kappa t}) \\ t\sigma_X^2 - \rho\frac{\sigma_X\sigma_S}{\kappa}(1 - e^{-\kappa t}) & \frac{\rho\sigma_X\sigma_S}{\kappa}(1 - e^{-\kappa t}) - \frac{\sigma^2}{2\kappa}(1 - e^{-2\kappa t}) & t\sigma_X^2 - 2\rho\frac{\sigma_X\sigma_S}{\kappa}(1 - e^{-\kappa t}) + \frac{\sigma_S^2}{2\kappa}(1 - e^{-2\kappa t}) \end{pmatrix}$$

and $\mathcal{N}(\mathbf{m}, \mathbf{V})$ denotes the normal distribution with mean vector \mathbf{m} and covariance matrix \mathbf{V} .

As we specify first the dynamics of the futures contract as a GBM and the logspread as an Ornstein-Uhlenbeck process the price of the futures contract leads the risk price. For example if the futures price is subject to a (demand or supply) shock the risk price follows and reduces the distance to the futures price. This asymmetric behavior is in line with empirical findings as there is strong evidence that futures prices lead the spot prices, e.g. see Chan (1987), Kawaller, Koch, and Koch (1987) and Stoll and Whaley (1990). Note that most studies investigate the relationship between a stock index and the corresponding futures contract. However, their main argument of less frequent trading in the spot market and differences in transaction costs are also valid in our setup. They both lead to asymmetric access to information which in turn results in an asymmetric behavior of the spread. Therefore, for the applications we have in mind, it seems natural to model the futures and the spread first, and then to derive the spot dynamics endogenously. Moreover, it is also possible to specify the relation in the reverse direction, i.e. to derive the dynamics of the futures price process based on the dynamics of the risk process

and the logspread. One can then proceed in a similar manner.

The model introduced has some similarities with Gaussian commodity spot models, e.g. with the ones discussed in Schwartz (1997) or with the more general model provided in Casassus and Collin-Dufresne (2005). In these models the triple of futures log price, spot log price and logspread is a 3-dimensional Gaussian process, too. These spot models, however, have different aims; e.g. they can be used for pricing long term forward commitments on the *same* commodity. Any forward position can be hedged by using one interest rate derivative and two short term futures contracts. This means that the latter models are *complete* and hence the model dynamics under the *risk-neutral measure* have to be calibrated to current futures and derivative prices.

The main aim of our model instead is to analyze the hedge error entailed when cross hedging risk exposures with futures written on a correlated, but *different* risk source. Our model includes a non-hedgeable risk factor, the spread, leading to *incompleteness*. We work under the physical measure since this is the only measure under which the hedge errors characteristics are relevant for risk management. Moreover, in a cross hedging situation a calibration is not always possible, e.g. if there are no liquid kerosene futures.

B. An Empirical Illustration

In the following we illustrate the estimation of the model by reconsidering the example of kerosene and crude oil. We use daily data of the spot kerosene price and the price of different crude oil futures contracts. The maturity of the futures contracts range from January 2009 until October 2010, resulting in 21 (overlapping) time series.

In a first step we check via the augmented Dickey-Fuller test whether the logspread of the kerosene spot price and the corresponding crude oil futures price is stationary. Table A1 in Appendix A reports the results for the different futures contracts. For most of the pairs we reject the null of a non-stationary logspread at any reasonable level. Of course, as one would expect

for a statistical test, the procedure does not suggest the existence of a long-term relationship for every pair even if it is present. For these cases the model uncertainty is obvious and we will check in Subsection A how the application of an optimal hedging strategy influences the hedging performance if the strategy is derived under our model but is applied to the 2GBM model, and vice versa.

Table A1 in Appendix A also presents the estimation result for our data sets. To concentrate on one asset in the remaining part of the paper we use one *representative contract* with moderate, not extreme, parameter values especially for κ and ρ . We choose the contract with maturity in August 2009. The number of observations at which both assets, the futures contract and the spot kerosene, are traded is 885. Figure 2 shows the time evolution of these two price series. Obviously, the price evolvments are very similar, which is also supported by the time-series plot of the logspread of the log prices (depicted in the lower panel).

In the next sections we derive the variance optimal hedge, the corresponding variance of the hedge error and derive quantitative and qualitative statements in terms of the structural parameters.

III. Optimal Variance Hedging with Futures Contracts

Suppose that a hedger sets up a portfolio consisting of futures contracts and cash positions, in order to hedge the risk position $h(I_T)$. In the following we denote by ξ_t the number of futures contracts held in the portfolio at time t . We assume that any futures position strategy $\xi = (\xi_t)_{t \in [0, T]}$ is non-anticipating, i.e. at any time it incorporates only information publicly available. In mathematical terms, this means that ξ is progressively measurable with respect to $(\mathcal{F}_t)_{t \geq 0}$, the filtration generated by the Brownian motions $(W^{(X)}, W^\perp)^\top$ and completed with the P -null sets of the basis.

If the futures price changes by ΔX_t from one trading day to the next, the hedger's margin

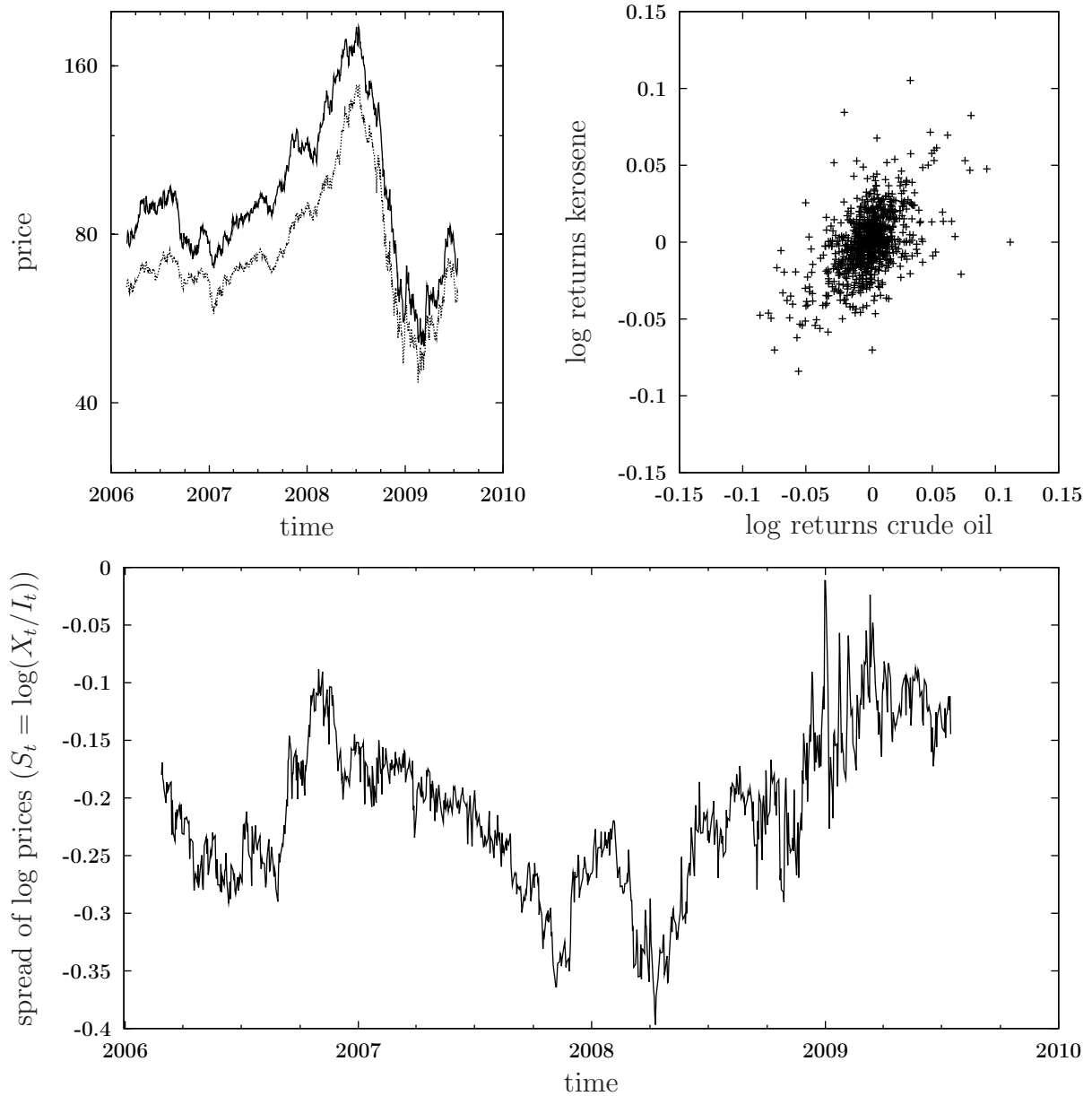


Figure 2: The upper left hand panel depicts the time evolution of the daily price of the crude oil futures with maturity in August 2009 in US\$/BBL (dashed line) and for spot jet kerosene in US\$/BBL (solid line) from 2006/02/27 until 2009/07/16 (resulting in 885 observations). The structure of this figure is the same as Figure 1 on page 3. The upper right hand panel exhibits the scatter plot of the daily returns. The lower panel depicts the time evolution of the logspread of the log prices. Note, that the units have been normalized just as in Figure 1.

account is adjusted by ΔX_t per futures contract. The cash position in the hedging portfolio is changed accordingly, entailing a portfolio value change due to variation margins of $\Delta V_t^{\text{mar}} = \xi_t \Delta X_t$.

Denote by $V = (V_t)_{t \in [0, T]}$ the *total value* of the hedging portfolio. Given an interest rate r , the cash position contributes to the portfolio by $rV_t dt$, hence the total value satisfies the continuous-time self-financing condition

$$(6) \quad dV_t = \xi_t dX_t + rV_t dt.$$

Equation (6) is linear. Therefore, given an initial portfolio value of $V_0 = v$, the portfolio process has the explicit representation

$$(7) \quad V_t = e^{rt} \left(v + \int_0^t e^{-rs} \xi_s dX_s \right),$$

(see e.g. Chapter 5.6 in Karatzas and Shreve (1991)).

Consider a self-financing hedge portfolio with futures position ξ_t at time t . The *conditional hedge error* of the portfolio at time $t \in [0, T]$ is then given by

$$(8) \quad \begin{aligned} C_t(\xi, v) &= \mathbb{E} \left(e^{-r(T-t)} h(I_T) | \mathcal{F}_t \right) - V_t \\ &= e^{rt} \left(\mathbb{E} \left(e^{-rT} h(I_T) | \mathcal{F}_t \right) - v - \int_0^t \xi_s e^{-rs} dX_s \right). \end{aligned}$$

$C_T(\xi, v)$ will also be referred to as the *realized hedge error*. Note that if $C_t(\xi, v)$ is negative, the combined value of the risk and the hedge portfolio is expected to end up with a plus.

To determine the variance optimal strategy within our model, i.e. the strategy minimizing the variance of the realized hedge error, we suppose that X is a martingale. This is a plausible assumption since X is a futures price process. Moreover, the empirical analysis of crude oil futures prices shows that the estimated drift parameter is close to zero for all contract months

and statistically insignificant for most assets (see Table A1 in Appendix A). In addition, as the estimation of the drift is notoriously challenging and very often highly speculative, it can easily distort the main aim of hedging, which is the reduction of risk. We therefore discuss here the martingale case in depth and postpone the discussion of the more general case to Section VI.

Assuming that X is a martingale means that $\mu = 0$ and $dX_t = \sigma_X X_t dW_t^{(X)}$. Then (8) implies that the discounted conditional hedge error is also a martingale. By applying a representation theorem from Stochastic Analysis, the martingale $e^{-rT} \mathbb{E}(h(I_T) | \mathcal{F}_t)$ can be written as a stochastic integral process of the form

$$(9) \quad e^{-rT} \mathbb{E}(h(I_T) | \mathcal{F}_t) = e^{-rT} \mathbb{E}(h(I_T)) + \int_0^t a_s dW_s^{(X)} + \int_0^t b_s dW_s^\perp, \quad t \in [0, T],$$

where a and b are progressively measurable and square-integrable processes (see e.g. Chapter 3.4 in Karatzas and Shreve (1991)). The first stochastic integral on the RHS is hedgeable, since it is driven by the same BM as the futures X . More precisely, following the strategy

$$(10) \quad \xi_t^* = \frac{a_t e^{rt}}{\sigma_X X_t},$$

the gain from the futures position up to time t satisfies $\int_0^t \xi_s^* e^{-rs} dX_s = \int_0^t a_s dW_s^{(X)}$. The second integral in (9) is orthogonal to $W^{(X)}$, and hence completely non-hedgeable with X . This implies that the strategy ξ^* minimizes the variance of the realized hedge error (see Theorem 1 in Föllmer and Sondermann (1986) where this argument has been employed for the first time). Moreover, the conditional hedge error satisfies

$$C_t(\xi^*, v) = e^{rt} \int_0^t b_s dW_s^\perp + \mathbb{E}(e^{-(T-t)r} h(I_T)) - e^{rt} v.$$

Profiting from the Markov property of the processes I and S , we may express a and b in terms of sensitivities of the expected risk with respect to the futures and the logspread values.

More precisely, let

$$\psi(t, x, s) = e^{-r(T-t)} \mathbb{E} \left(h \left(X_T^{t,x} e^{-S_T^{t,s}} \right) \right),$$

where $X^{t,x}$ and $S^{t,s}$ are the solutions of (1) resp. (2) on $[t, T]$ with initial conditions $X_t^{t,x} = x$ resp. $S_t^{t,s} = s$. We will refer to ψ as the *value function*. If h is Lipschitz continuous and its weak derivative h' is Lebesgue-almost everywhere differentiable, then ψ is continuously differentiable with respect to x and s , and

$$(11) \quad \psi_x(t, x, s) = \frac{\partial}{\partial x} \psi(t, x, s) = e^{-r(T-t)} \mathbb{E} \left(h' \left(X_T^{t,x} e^{-S_T^{t,s}} \right) X_T^{t,x} e^{-S_T^{t,s}} \right).$$

For details we refer the interested reader to Lemma 4.8. in Ankirchner and Heyne (2011), where a similar statement is shown. Notice that $\partial S_T^{t,s} / \partial s = e^{-\kappa(T-t)}$, and hence by the same reasoning

$$(12) \quad \begin{aligned} \psi_s(t, x, s) &= \frac{\partial}{\partial s} \psi(t, x, s) = -e^{-r(T-t)} \mathbb{E} \left(h' \left(X_T^{t,x} e^{-S_T^{t,s}} \right) X_T^{t,x} e^{-S_T^{t,s}} e^{-\kappa(T-t)} \right) \\ &= -e^{-\kappa(T-t)} x \frac{\partial}{\partial x} \psi(t, x, s). \end{aligned}$$

The pair $(X, S)^\top$ is a 2-dimensional SDE, driven by $(W^{(X)}, W^\perp)^\top$ via the diffusion matrix

$$\tilde{\Sigma}(x, s) = \begin{pmatrix} \sigma_X x & 0 \\ \rho \sigma_S & \bar{\rho} \sigma_S \end{pmatrix}.$$

With Itô's formula we obtain that the processes a and b , appearing in the martingale representation (9), are given by

$$\begin{pmatrix} a_t \\ b_t \end{pmatrix} = e^{-rt} \tilde{\Sigma}^\top(X_t, S_t) \begin{pmatrix} \psi_x(t, X_t, S_t) \\ \psi_s(t, X_t, S_t) \end{pmatrix}.$$

From (10) and (12) we can deduce the following result describing the variance optimal hedge in terms of the futures-delta, and the minimal hedge error in terms of the logspread-delta.

Theorem III.1. *The variance optimal futures position in the hedging portfolio is given by*

$$(13) \quad \xi_t^* = \left[1 - \frac{\sigma_S}{\sigma_X} \rho e^{-\kappa(T-t)} \right] \psi_x(t, X_t, S_t),$$

and entails a realized hedge error of

$$(14) \quad C_T(\xi^*, v) = \mathbb{E}(h(I_T)) - e^{rT} \left(v - \bar{\rho} \int_0^T e^{-rt} \sigma_S \psi_s(t, X_t, S_t) dW_t^\perp \right).$$

Observe that the optimal hedge ξ^* is the Delta of the position's expectation, dampened by the *hedge ratio* defined by

$$f(T-t) = 1 - \frac{\sigma_S}{\sigma_X} \rho e^{-\kappa(T-t)}.$$

The factor essentially equals 1 if the product of time to maturity $T-t$ and reversion speed κ is large. In this case the logspread is expected to return to its mean reversion level before maturity, and hence the position should be fully hedged with the futures. As maturity approaches, the short-term fluctuations have an increasing impact on the terminal hedge performance, making the hedge ratio converge to

$$(15) \quad h = \rho_{IX} \frac{\sigma_I}{\sigma_X}.$$

Indeed, due to (3), we have $\lim_{(T-t) \downarrow 0} 1 - \sigma_S \rho e^{-\kappa(T-t)} / \sigma_X = 1 - \sigma_S \rho / \sigma_X = h$. We remark that h , defined in (15), is sometimes referred to as the *minimum variance hedge ratio* (see e.g. Chapter 3.4. in Hull (2008)).

Observe that if κ is equal to zero, which essentially means that there is no mean reversion,

then the logspread is not stationary and its variance increases linearly with time. The hedge ratio is not dampened and it coincides with h . In this case the strategy ξ^* is equal to the optimal strategy in a model where both X and I are modeled as GBMs (see Section A for more details).

In Formula (13) the optimal hedge is expressed in terms of the Delta with respect to the futures price. In order to obtain a representation in terms of the Delta with respect to the illiquid asset price I , define first $\varphi(t, y, s) = e^{-r(T-t)}\mathbb{E}(h(I_T^{t,y,s}))$, where $I^{t,y,s}$ is the solution of the SDE for the illiquid asset on $[t, T]$, with initial values $I_t^{t,y,s} = y$ and $S_t^{t,s} = s$. Note that $\psi(t, x, s) = \varphi(t, e^{-s}x, s)$, and in particular, $\psi_x(t, x, s) = e^{-s}\varphi_y(t, e^{-s}x, s)$. Thus, the optimal hedge may be rewritten as

$$(16) \quad \xi_t^* = e^{-S_t} \left[1 - \frac{\sigma_S}{\sigma_X} \rho e^{-\kappa(T-t)} \right] \varphi_y(t, I_t, S_t).$$

If the logspread is positive, then the illiquid asset price is expected to rise relative to the futures price. This explains why in Equation (16) the delta is reduced by the factor e^{-S_t} . Conversely, if the logspread is negative, then the illiquid asset is expected to fall relative to the futures price. In this case the delta is augmented by the factor e^{-S_t} .

Finally, we remark that the hedge ratio remains the same if the hedger uses an option on the futures for hedging the risk exposure $h(I_T)$. Denote by $\Delta(t, x)$ the option's delta at time t , given a futures price of $X_t = x$. The dynamics of the option price $P(t, X_t)$ satisfy $dP(t, X_t) = rP(t, X_t) dt + \Delta(t, X_t) dX_t$, and the value of a self-financing portfolio containing ξ_t options at time t is given by

$$(17) \quad V_t = e^{rt} \left(V_0 + \int_0^t e^{-rs} \xi_s \Delta(s, X_s) dX_s \right).$$

The variance minimizing *option position* can be shown to be equal to

$$(18) \quad \xi_t^* = \left[1 - \frac{\sigma_S}{\sigma_X} \rho e^{-\kappa(T-t)} \right] \frac{\psi_x(t, X_t, S_t)}{\Delta(t, X_t)}.$$

Suppose that a non-linear position of kerosene is hedged with an option, written on a crude oil futures, having a similar payoff profile. Then the ratio of deltas $\psi_x(t, X_t, S_t)/\Delta(t, X_t)$ is usually stable and hence the hedging portfolio does not need to be rebalanced as strongly as when using futures for hedging.

IV. Standard Deviation of the Hedge Error

Having derived the variance optimal strategy and the corresponding hedge error, see Theorem III.1, we now aim at computing the implied standard deviation of the hedge error. This allows us to quantify the risk associated with the optimal strategy, which is important for risk management and performance evaluation of the hedging strategy. We therefore derive analytic and semianalytic formulas for the standard deviation of the hedge error when minimizing the variance of risk exposures within our model. As in the previous section, we assume that the hedger does not have any directional view concerning the futures. This means that the futures price X is a martingale with dynamics $dX_t = \sigma_X X_t dW_t$.

We aim at computing the standard deviation of the hedge error when cross-hedging the position $h(I_T)$ following the strategy ξ^* of (13). Note that the standard deviation of $C_T(\xi^*, v)$ coincides with the standard deviation of $e^{rT} \bar{\rho} \int_0^T e^{-rt} \sigma_S \psi_s(t, X_t, S_t) dW_t^\perp$. The Itô isometry for stochastic integrals implies

$$(19) \quad \text{std}(C_T(\xi^*, v)) = e^{rT} \bar{\rho} \sigma_S \sqrt{\int_0^T e^{-2rt} \mathbb{E}(\psi_s^2(t, X_t, S_t)) dt}.$$

In general, there is no closed-form expression for the formula for the integral in (19). For linear

positions, however, we may explicitly calculate the variance, since the futures and the spread are lognormally distributed. Besides, for positions corresponding to Plain Vanilla options, the logspread-delta ψ_s has an explicit representation, and thus allows for an efficient Monte Carlo simulation of the error (19). We proceed by discussing both cases separately.

A. Linear Positions

In this subsection we derive an analytic formula for the hedge error variance when cross-hedging a linear position. This is the most relevant case, since most of the risk positions of industrial companies are linear. Think for instance of an airline being exposed to a short position of kerosene.

Suppose that the payoff function h given by $h(y) = cy$, with $c \in \mathbb{R}$. In this case the delta of the value function with respect to the futures price satisfies $\psi_x(t, x, s) = e^{-r(T-t)} \mathbb{E} \left(cX_T^{t,1} e^{-S_T^{t,s}} \right)$ (see (11)). Thus, with (12), we get

$$\frac{\partial}{\partial s} \psi(t, x, s) = -e^{-\kappa(T-t)} x e^{-r(T-t)} \mathbb{E} \left(cX_T^{t,1} e^{-S_T^{t,s}} \right).$$

In the following we do not only need to compute the expectation of the product $X_T^{t,1} e^{-S_T^{t,s}}$, but also the expectation of the product of higher moments of the logspread and the illiquid asset. We therefore straightly provide the following lemma.

Lemma IV.1. *Let $a \in \mathbb{R}$ and $b \in \mathbb{R}_+$, then*

$$\mathbb{E} \left(e^{-aS_t^{0,s}} (X_t^{0,x})^b \right) = A(a, b, x, s, t),$$

where $A(a, b, x, s, t)$ is defined by

$$(20) \quad A(a, b, x, s, t) = x^b \exp \left[-\frac{1}{2} \sigma_X^2 t (b - b^2) - a s e^{-\kappa t} - a \left(m + b \rho \sigma_X \sigma_S \frac{1}{\kappa} \right) (1 - e^{-\kappa t}) + \frac{1}{2} a^2 \sigma_S^2 \frac{1}{2\kappa} (1 - e^{-2\kappa t}) \right].$$

With this at hand the standard deviation (19) simplifies to

$$(21) \quad \text{std}(C_T(\xi^*, v)) = |c| \bar{\rho} \sigma_S \sqrt{\int_0^T e^{-2\kappa(T-t)} \mathbb{E}(X_t^2 A^2(1, 1, 1, S_t, T-t)) dt}.$$

From this we are able to derive the following explicit formula for the hedge error variance.

Theorem IV.2. *The variance optimal cross-hedge of a linear position cI_T entails a hedge error with standard deviation*

$$(22) \quad \begin{aligned} \text{std}(C_T) &= \sigma_S \sqrt{1 - \rho^2} x \exp \left((m - s) e^{-\kappa T} - m - \rho \sigma_X \sigma_S \frac{1}{\kappa} (1 - 2e^{-\kappa T}) + \sigma_S^2 \frac{1}{4\kappa} (1 - 2e^{-\kappa T}) \right) \\ &\quad \times |c| \sqrt{\int_0^T \exp \left(-2\kappa(T-t) + \sigma_X^2 t - 2\rho \sigma_X \sigma_S \frac{1}{\kappa} e^{-\kappa(T-t)} + \sigma_S^2 \frac{1}{2\kappa} e^{-2\kappa(T-t)} \right) dt}. \end{aligned}$$

The integral in (22) can be computed in a straightforward manner using standard numerical quadratures algorithms.

When analyzing the dependence of the hedge error on the different model parameters it is convenient to rewrite the hedge error formula (22) as follows:

$$(23) \quad \text{std}(C_T) = |c| \sigma_S \sqrt{1 - \rho^2} x \exp \left((m - s) e^{-\kappa T} - m \right) \left[\int_0^T \exp \left(-2\kappa(T-t) + \sigma_X^2 t \right) \times \exp \left(-2\rho \sigma_X \frac{\sigma_S}{\kappa} (1 + e^{-\kappa(T-t)} - 2e^{-\kappa T}) + \frac{\sigma_S^2}{2\kappa} (1 + e^{-2\kappa(T-t)} - 2e^{-\kappa T}) \right) dt \right]^{\frac{1}{2}}.$$

The hedge error (22) can be approximated with simplified formulas. For long maturities, i.e. large T , the factor $\exp((m-s)e^{-\kappa T} - m)$ approximately coincides with e^{-m} . Moreover,

$$\int_0^T e^{-2\kappa(T-t) + \sigma_X^2 t} dt = \frac{e^{\sigma_X^2 T} - e^{-2\kappa T}}{2\kappa + \sigma_X^2} \approx \frac{e^{\sigma_X^2 T}}{2\kappa + \sigma_X^2},$$

and hence

$$\text{for long maturities:} \quad \text{std}(C_T) \approx |c| x e^{-m} \sigma_S \sqrt{1 - \rho^2} e^{\left(-\rho \sigma_S \sigma_X + \frac{\sigma_S^2}{4}\right)/\kappa} \sqrt{\frac{e^{\sigma_X^2 T}}{2\kappa + \sigma_X^2}}.$$

Observe that the hedge error increases exponentially with time to maturity. However, the volatility squared σ_X^2 is usually low (see Table A1 in Appendix A), and thus the hedge error increases approximately linearly, with slope $\sigma_X^2/2$, in the first several years (compare with the upper left panel in Figure 3).

For short maturities, i.e. small T , the factor $\exp((m-s)e^{-\kappa T} - m)$ approximately coincides with e^{-s} . Besides, by linearly approximating exponentials with the Taylor expansion up to the first order, we have

$$\int_0^T e^{-2\kappa(T-t) + \sigma_X^2 t} dt = \frac{e^{-\sigma_X^2 T} - e^{-2\kappa T}}{2\kappa + \sigma_X^2} \approx T.$$

Therefore, we get the approximation

$$\text{for short maturities:} \quad \text{std}(C_T) \approx |c| \sigma_S \sqrt{1 - \rho^2} x e^{-s} \sqrt{T}.$$

Note that the hedge error increases with order $T^{1/2}$, for short maturities (see the upper left panel in Figure 3). The kink in Figure 3 is determined by how fast the Ornstein Uhlenbeck process describing the logspread attains its stationary distribution. The variance of the logspread at time T , as a function of time to maturity $T - t$, is given by $\sigma_S^2(1 - e^{-2\kappa(T-t)})/(2\kappa)$. The hedge

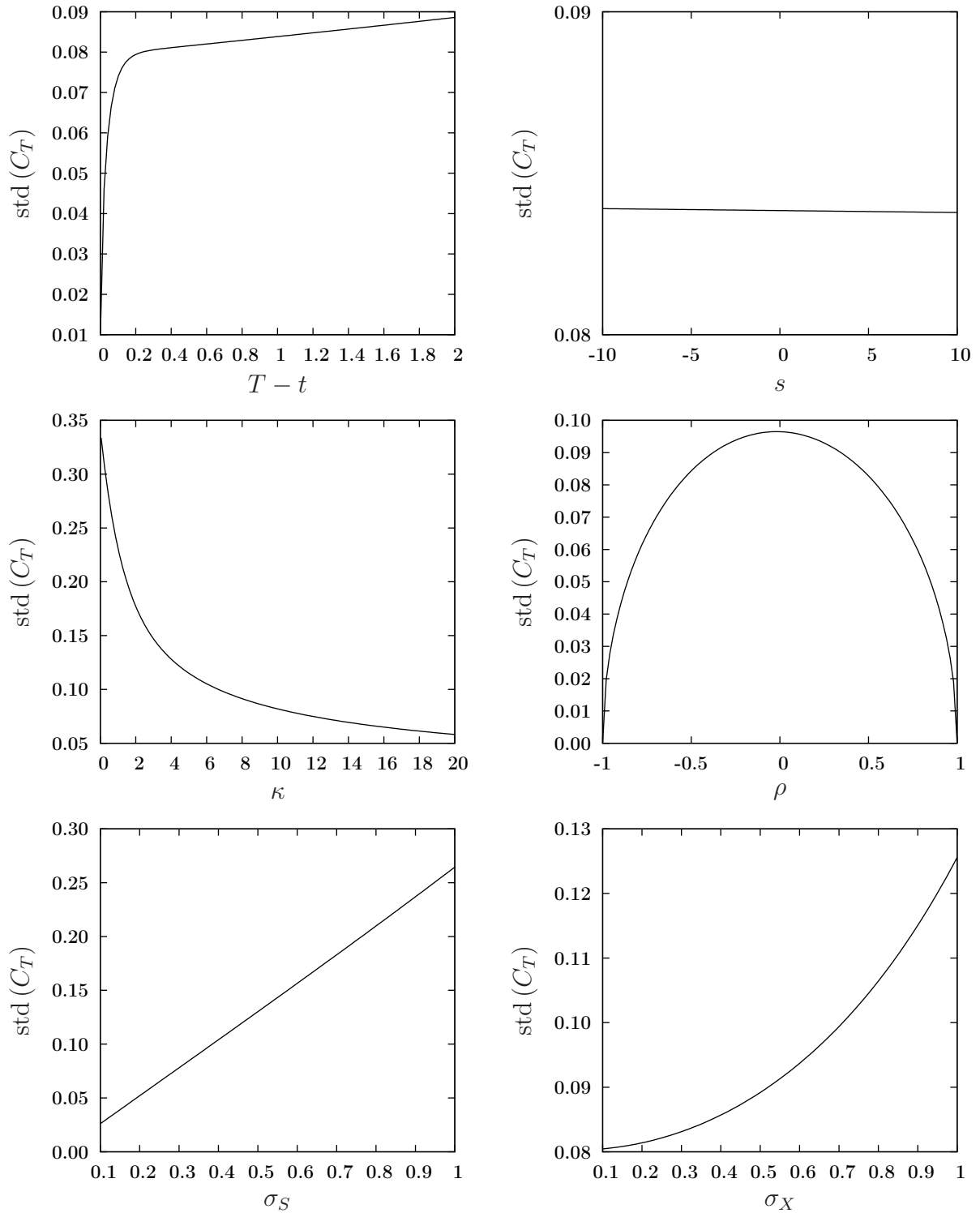


Figure 3: This figure shows the sensitivity of the standard deviation of the hedge error with respect to the time to maturity (in the upper left hand panel) and with respect to the parameters of the model (remaining panels). In each panel only the parameter indicated on the abscissa is varied while the others remain fixed at the estimates from the futures contract with maturity in August 2009.

error is roughly proportional to the variance of the logspread.

The hedge error vanishes as the variance of the stationary logspread distribution, $\sigma_S^2/(2\kappa)$, converges to zero. Moreover, it is straightforward to show

$$(24) \quad \lim_{\kappa \downarrow 0} \text{std}(C_T) = |c| x e^{-s} \frac{\sigma_S}{\sigma_X} \sqrt{1 - \rho^2} \sqrt{e^{\sigma_X^2 T} - 1}.$$

Of course, not only the impact of the time to maturity to the hedge error is of interest but also the influence of the model's parameters. Figure 3 highlights the sensitivity of the standard deviation of the hedge error towards changes in the parameters. The figure depicts the resulting standard deviation by changing one parameter and keeping the others constant. The fixed parameters are set to the estimates of the futures contract with maturity in August 2009. The figure shows (in the middle left panel) the decreasing standard deviation in the mean reversion speed κ . This comes at no surprise as a larger κ results in a faster return to the long-term relationship. The reverse U-shaped behavior with respect to the correlation ρ (in the right middle panel) highlights the change from the incomplete market setting for $|\rho| < 1$ to the complete market setting for $|\rho| = 1$. With increasing instantaneous variance of the logspread and the futures price process (σ_S and σ_X) the variance of the hedge error also increases (shown in the lower panels).

B. Non-linear Positions

In practice the case of non-linear risk positions is also relevant. For instance, consider the very illiquid German natural gas futures markets. Due to the illiquidity gas traders frequently use futures of neighboring countries for hedging purposes. So, if operators of German gas power plants protect themselves against changing gas prices by buying Dutch gas on the futures market (e.g. natural gas futures of the Dutch market TTF) the basis risk is due to a geographical spread in commodity prices which arises from different trading places for the same underlying. In this

case TTF contracts serve as proxies that are cointegrated with the natural gas prices in the German market area. The profit margin of a gas power plant is essentially determined by the spark spread, i.e. the spread between the electricity price per MWh and the price of the amount of gas the plant needs for producing 1 MWh of electricity. Electricity will only be produced if the profit margin exceeds the operating costs. A gas power plant can thus be seen as a call option, a highly non-linear position, on the spark spread. We therefore also consider here the hedging of non-linear risk positions.

When it comes to hedging non-linear risk positions, the problem is that in general there are no explicit formulas available for the standard deviation of the hedge error. We explain here how a swift simulation analysis can be used to obtain the hedge error characteristics for standard options.

The simulation-based hedge error analysis basically works as follows. First simulate N independent paths of the futures price and the logspread. Then calculate the performance of the hedging portfolio along any simulation path and subtract it from the risk position $h(I_T)$. The collection of hedge errors obtained in this way may be analyzed with respect to the empirical standard deviation, median, and other statistical characteristics.

In order to calculate the portfolio value along any simulation path, one needs to compute the portfolio position and hence the delta at *any* time discretization point. As we will show below, for Plain Vanilla options there are analytic formulas for the deltas, allowing thus to quickly calculate the portfolio performance for every simulation path. Indeed, the deltas resemble the deltas for Plain Vanilla options in the Black-Scholes model. As an example we provide the relevant formulas for a European call option with strike K , i.e. h is given by $h(y) = (y - K)^+$.

First observe that the value function ψ for call options is given by

$$\psi(t, x, s) = \mathbb{E} \left((I_T^{t,x,s} - K)^+ \right).$$

From the proof of Lemma IV.1 it can be seen that $I_T^{t,x,s} = e^{-S_T^{t,s}} X_T^{t,x}$ can be written as

$$I_T^{t,x,s} = x \exp \left(-\frac{1}{2} \sigma_X^2 (T-t) - s e^{-\kappa(T-t)} - m(1 - e^{-\kappa(T-t)}) \right) \exp(\sigma N),$$

where N is a standard normal variable and σ^2 is given by

$$\sigma^2 = \sigma_X^2 (T-t) - 2\rho\sigma_X\sigma_S \frac{1}{\kappa} (1 - e^{-\kappa(T-t)}) + \sigma_S^2 \frac{1}{2\kappa} (1 - e^{-2\kappa(T-t)}).$$

We get

$$(25) \quad \psi(t, x, s) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \left(x e^{-\frac{1}{2} \sigma_X^2 (T-t) - s e^{-\kappa(T-t)} - m(1 - e^{-\kappa(T-t)}) + \sigma y} - K \right)^+ e^{-y/2} dy.$$

In analogy to the standard Black-Scholes case we define the functions $d_+(t, x, s)$ and $d_-(t, x, s)$ as

$$d_+(t, x, s) = \frac{1}{\sigma} \left[\log \left(\frac{K}{x} \right) + \frac{1}{2} \sigma_X^2 (T-t) + s e^{-\kappa(T-t)} + m(1 - e^{-\kappa(T-t)}) \right]$$

and $d_-(t, x, s) = d_+(t, x, s) - \sigma$.

Note that the integrand in the integral in (25) equals zero if $y < d_+(t, x, s)$, and hence

$$\begin{aligned} \psi(t, x, s) &= x \exp \left(-\frac{1}{2} \sigma_X^2 (T-t) - s e^{-\kappa(T-t)} - m(1 - e^{-\kappa(T-t)}) + \frac{1}{2} \sigma^2 \right) \Phi(d_-(t, x, s)) \\ &\quad - K \Phi(d_+(t, x, s)), \end{aligned}$$

where Φ denotes the cumulative distribution function of the standard normal distribution. The above explicit representation of $\psi(t, x, s)$ yields

$$\begin{aligned} \psi_x(t, x, s) &= \exp \left(-\frac{1}{2} \sigma_X^2 (T-t) - s e^{-\kappa(T-t)} - m(1 - e^{-\kappa(T-t)}) + \frac{1}{2} \sigma^2 \right) \\ &\quad \times [\Phi(d_-(t, x, s)) + \varphi(d_-(t, x, s)) \partial_x d_-(t, x, s)] - K \varphi(d_+(t, x, s)) \partial_x d_+(t, x, s), \end{aligned}$$

where $\partial_x d_+(t, x, s) = \partial_x d_-(t, x, s) = 1/(\sigma x)$.

The analytic and semianalytic formulas for the hedge error allow the efficient computation and the comparison of the hedge error variance for different, potentially useable, liquid futures contracts. Up to now, however, we assume that we hedge within the correct model (with a stationary logspread). Although, statistical tests may help to decide whether the logspread is stationary or not, there is always the risk to *hedge within the wrong model* and a relevant question arises: how sensitive is the hedge error with respect to the model choice? We address this question in the next Section.

V. The Performance of Suboptimal Hedging Strategies

So far we have assumed that we know with certainty that the price of the illiquid asset and the price of the liquid futures contract are cointegrated and evolve according to our model. However, it may happen that a statistical test leads to a wrong conclusion or different tests lead to different implications. In other words, we face model uncertainty.

In the following Subsection A we consider a 2GBM model and derive the hedge error obtained by using the optimal strategy from our model. Furthermore, we analyse the impact of applying the optimal hedging strategy from the 2GBM model to our model. We then proceed by comparing the optimal *dynamic* hedge with its optimal *static* counterpart. In practice, traders often hedge linear positions statically, holding a position in futures that corresponds to the size of the risk. By fully hedging the risk they intuitively reflect that the hedge ratio essentially equals 1 whenever time to maturity is long. In Subsection B we first derive the optimal *static* hedging strategy and compare it with the hedging strategy ξ^* , which allows for portfolio regrouping.

A. The Costs of Ignoring a Long-term Relationship or Falsely Assuming a Long-term Relationship

We next introduce a simple model where both X and I are GBMs, hence are not cointegrated and the logspread does not have a stationary distribution. We will refer to this model as the 2GBM model.

In both models, the futures price is assumed to satisfy the dynamics

$$dX_t = \sigma_X X_t dW_t^{(X)},$$

but in contrast to the model with a stationary logspread, discussed in the previous sections, the 2GBM model assumes that the illiquid asset price process is also a GBM with dynamics

$$dI_t = \sigma_I I_t (\rho_{IX} dW_t^{(X)} + \bar{\rho}_{IX} dW_t^\perp).$$

In this model the variance minimizing hedging strategy for European options with payoff function h are known to have the simple form

$$(26) \quad \zeta_t = \rho_{IX} \frac{\sigma_I I_t}{\sigma_X X_t} \psi_y(t, I_t),$$

where $\psi(t, y) = e^{-r(T-t)} \mathbb{E}(h(I_T^{t,y}))$. For a derivation of (26) we refer to Hulley and McWalter (2008); see also Ankirchner and Heyne (2011) for a derivation in a slightly more general setting using Backward Stochastic Differential Equations (BSDEs).

The optimal cross-hedge within the 2GBM models (26), is essentially determined by the cross correlation. If an airline company used a 2GBM model estimated with daily data to hedge kerosene short positions, then it would considerably underhedge its kerosene risk, facing thus unnecessarily high variations in costs. But, by how much does the hedge error increase if

we use the wrong model? We next quantify the risk by calculating the hedge error when using the optimal strategy ζ of the 2GBM model while the log prices are cointegrated and evolve according to the dynamics of our model.

We restrict our analysis to linear positions of the form cI_T . As before we denote the realized hedge error by

$$C_T = cI_T - e^{rT} \left(v + \int_0^T e^{-rs} \zeta_s dX_s \right).$$

The following proposition provides the hedge error variance for the strategy (26) under our model with cointegration and under the 2GBM model.

Proposition V.1. *Hedging the linear position cI_T with the strategy ζ entails a hedge error in the cointegration model with variance*

$$\begin{aligned} \mathbb{V}(C_T) = c^2 & \left\{ A(2, 2, X_0, S_0, T) - A^2(1, 1, X_0, S_0, T) - 2 \int_0^T \rho_{IX} \sigma_I (\sigma_X - \rho \sigma_S e^{-\kappa(T-t)}) B_t dt \right. \\ & \left. + \int_0^T \rho_{IX}^2 \sigma_I^2 A(2, 2, X_0, S_0, t) dt \right\}, \end{aligned}$$

where B_t is given by

$$\begin{aligned} B_t = X_0^2 & \exp \left(\sigma_X^2 t - \frac{1}{\kappa} \sigma_S \sigma_X \rho (3 - 2e^{-\kappa T} - 2e^{-\kappa t} + e^{-\kappa(T-t)}) - S_0 (e^{-\kappa T} + e^{-\kappa t}) \right) \\ & \times \exp \left(\frac{1}{4\kappa} \sigma_S^2 (2 - e^{-2\kappa T} - e^{-2\kappa t} + 2e^{-\kappa(T-t)} - 2e^{-\kappa(T+t)}) - m(2 - e^{-\kappa T} - e^{-\kappa t}) \right). \end{aligned}$$

Under the correct model, the 2GBM model, the minimal variance of the realized hedge error is given by

$$(27) \quad \mathbb{V}(C_T) = c^2 y^2 (1 - \rho_{IX}^2) \left(e^{\sigma_I^2 T} - 1 \right).$$

Proposition V.1 allows us to analyse the ignorance of a long-term relationship with respect to the variance of the hedge error. Of course the variance of the optimal strategy ξ^* given by Equation (13) is less than the standard error using the strategy ζ from the 2GBM model given by Equation (26). The upper panels of Figure 4 compare the performance of the two strategies.³ When following the strategy ζ , the risk position is underhedged, yielding the hedge error (dashed line) to grow continually with time to maturity. The hedge error does not flatten as strongly as when following strategy ξ^* , whose corresponding hedge error is depicted by the solid line. For very short maturities, the mean reversion has little time to develop and hence the hedge error entailed by ζ is similar to the hedge error of ξ^* . For long maturities, however, ζ is considerably outperformed by ξ^* , leading for example to a more than three times higher error standard deviation over a two year hedging period.

Since by definition there is no strategy with a smaller variance than the variance optimal strategy the proportion of a three times larger value is strongly convincing. To fairly compare our model with the 2GBM model, we also consider the inverse case, that is we study the hedge error of the optimal strategy from the model with the stationary logspread, ξ^* , when there is no cointegration. To this end we have to derive the resulting hedge error in the 2GBM model which is given in the next proposition.

Proposition V.2. *Hedging the linear position cI_T with the strategy ξ^* , see (13), entails a hedge*

³Note that we used the estimated parameters obtained for the August 2009 crude oil futures contract (see Table A1) and that the correlation ρ_{IX} can be expressed in terms of the structural parameters of our model, see (4).

error in the 2GBM model with variance

$$(28) \quad \mathbb{V}(C_T) = c^2 \left\{ B(0, 2, T) - B^2(0, 1, T) \right. \\ \left. - 2 \int_0^T \sigma_I \rho \left[1 - \frac{\sigma_S}{\sigma_X} \rho e^{-\kappa(T-t)} \right] A(1, 1, 1, 0, T-t) \sigma_X B(1 - e^{-\kappa(T-t)}, 1 + e^{-\kappa(T-t)}, t) dt \right. \\ \left. + \int_0^T \left[1 - \frac{\sigma_S}{\sigma_X} \rho e^{-\kappa(T-t)} \right]^2 A^2(1, 1, 1, 0, T-t) \sigma_X^2 B(2 - 2e^{-\kappa(T-t)}, 2e^{-\kappa(T-t)}, t) dt \right\},$$

where $B(a, b, t)$ is given by

$$(29) \quad B(a, b, t) = \mathbb{E}(X_t^a I_t^b) = X_0^a I_0^b \exp\left(\frac{1}{2}t [\sigma_X^2(a^2 - a) + 2ab\sigma_X\sigma_I\rho + \sigma_I^2(b^2 - b)]\right).$$

Note that now not all parameters are identified. This is in contrast to the previous scenario, where we investigated the impact of ignoring a long-term relationship. As the logspread is not stationary in the 2GBM model the parameter κ is implicitly set to zero. However, to provide a realistic comparison we estimate the implied distribution of $\hat{\kappa}$ in the following way: we simulate 10000 sample paths from the 2GBM model with $T = 885$ observations. Based on these time series we estimate our model leading to the distribution of $\hat{\kappa}$. For the graphical illustration we use the corresponding 10%, 50% and 90% quantiles leading to 0.0927, 0.8365 and 2.2450 respectively. The dashed lines in the lower left panel of Figure 4 show the hedge error standard deviation, for these different mean reversion speeds, when the real prices behave as in the 2GBM model, but the risk is hedged according to the cointegration model optimal strategy ξ^* . As a benchmark, the panel depicts also the genuine minimal error standard deviation (solid line) implied by the optimal strategy ζ . The smaller the mean reversion speed, the smaller the hedge error. Moreover, the hedge error converges to the minimal hedge error as the mean reversion speed converges to zero, showing that the cointegration model embeds the 2GBM model.

In many real world applications, it may not be obvious that there is a long-term relationship between the hedging instrument and the risk to be hedged. Comparing the left upper and left

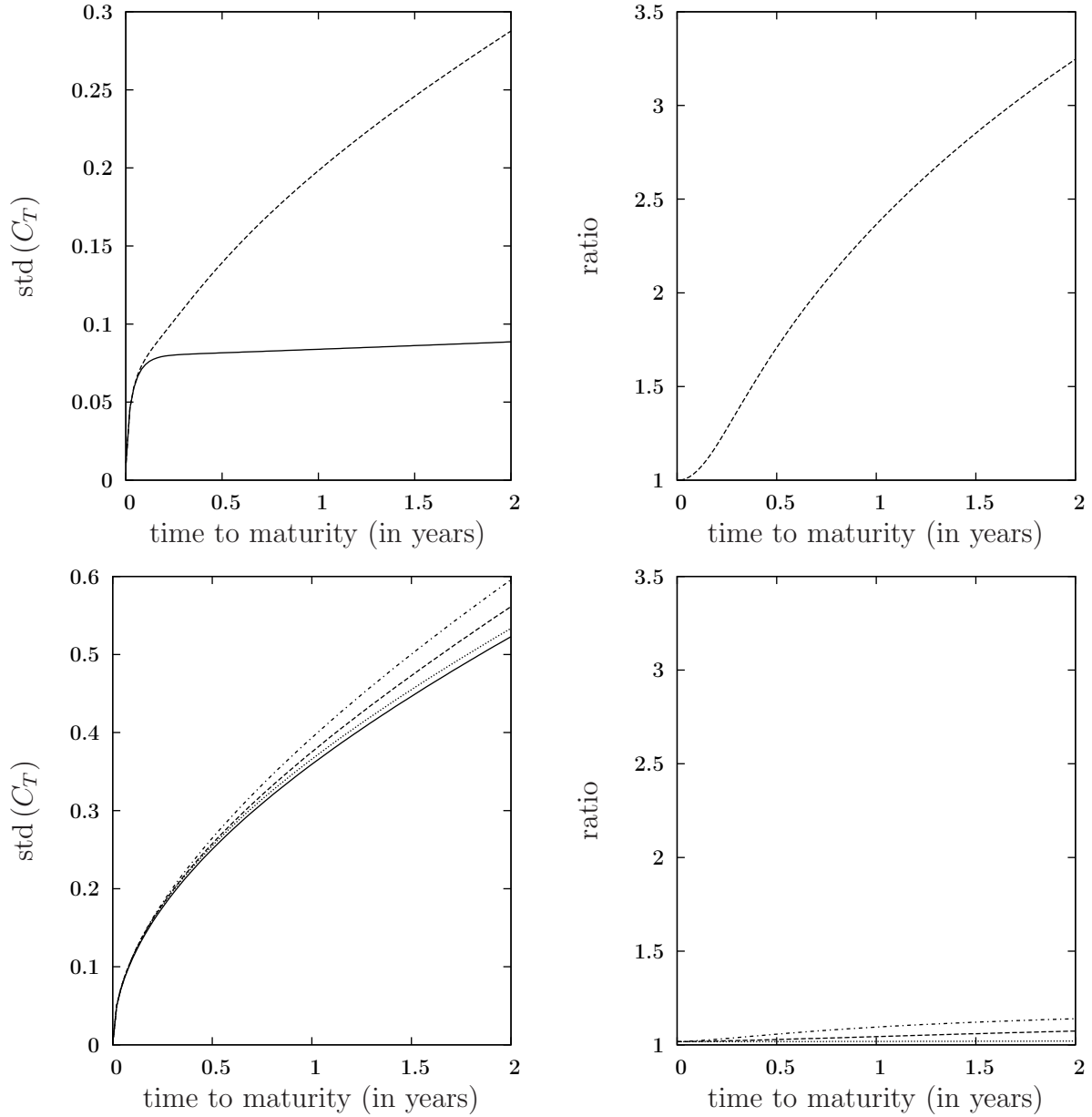


Figure 4: The upper left panel shows the standard deviation of the hedge error under cointegration using the optimal strategy (solid line) and the strategy from the 2GBM model (dashed line). The lower left panel shows the standard deviation under the 2GBM model using the optimal strategy (solid line) and the strategy from the cointegrated model (dashed lines) for $\kappa \in \{2.2450, 0.8365, 0.0927\}$, from top to down, respectively. The panels on the right depict the ratio of the standard deviation of the strategies.

lower panel of Figure 4 we conclude that in ambiguous situations it is nevertheless better to use a model allowing for a stationary spread rather than using a simpler model that does not. The error implied by mistakenly assuming a stationary logspread is significantly smaller than the error made by mistakenly assuming that the hedging instrument is not cointegrated with the risk.

B. The Costs of Using a Static Hedge

In practice linear positions are often only statically hedged, even though the variance minimizing hedge is not constant. In the numerical example below we compare the standard deviations of static and dynamic hedges of a linear position, and address the question by how much the *dynamic* variance minimizing strategy outperforms the static one.

For that purpose we need to derive the optimal static hedge position $a \in \mathbb{R}$ that minimizes the variance

$$C_T(a) = cI_T - e^{rT} \left(v + \int_0^T e^{-rt} a \, dX_t \right).$$

With this at hand we can calculate the minimal error standard deviation that can be achieved by hedging statically with futures.

Proposition V.3. *The optimal static hedging position \check{a} in futures contracts which minimizes the variance of the hedge error is given by*

$$(30) \quad \check{a} = \frac{\text{Cov} \left(cI_T, e^{rT} \int_0^T e^{-rt} \, dX_t \right)}{\mathbb{V} \left(e^{rT} \int_0^T e^{-rt} \, dX_t \right)} = ce^{-rT} \frac{\mathbb{E} \left(I_T \int_0^T e^{-rt} \, dX_t \right)}{\sigma_X^2 \mathbb{E} \left(\int_0^T e^{-2rt} X_t^2 \, dt \right)}$$

with corresponding variance

$$(31) \quad \mathbb{V}(C_T(\check{a})) = \mathbb{E}(C_T^2(\check{a})) = c^2 \mathbb{V}(I_T) - \frac{c^2 \left[\mathbb{E} \left(I_T \int_0^T e^{-rt} dX_t \right) \right]^2}{\sigma_X^2 \mathbb{E} \left(\int_0^T e^{-2rt} X_t^2 dt \right)}.$$

The expectation in the denominator is given by

$$\mathbb{E} \left(\int_0^T e^{-2rt} X_t^2 dt \right) = \begin{cases} \frac{X_0^2}{\sigma_X^2 - 2r} (e^{(\sigma_X^2 - 2r)T} - 1), & \text{if } \sigma_X^2 \neq 2r, \\ X_0^2 T, & \text{if } \sigma_X^2 = 2r. \end{cases}$$

For the expectation in the numerator we have, assuming $\sigma_X^2 > r$,

$$(32) \quad \mathbb{E} \left(I_T \int_0^T e^{-rt} dX_t \right) = \begin{cases} X_0^2 e^{\lambda(T)} \frac{\sigma_X^2}{\sigma_X^2 - r} (e^{(\sigma_X^2 - r)T} - 1), & \text{if } \rho = 0, \\ X_0^2 e^{\lambda(T) + \frac{\rho \sigma_X \sigma_S}{\kappa} e^{-\kappa T}} (\Lambda_1(T) - \Lambda_2(T)), & \text{if } \rho \neq 0, \end{cases}$$

with

$$\begin{aligned} \Lambda_1(T) &= \frac{\sigma_X^2 e^{(\sigma_X^2 - r)T} (|\rho| \sigma_X \sigma_S)^{-\frac{\sigma_X^2 - r}{\kappa}}}{\kappa^{1 - \frac{\sigma_X^2 - r}{\kappa}}} \left(\gamma \left(\frac{\sigma_X^2 - r}{\kappa}, \frac{1}{\kappa} |\rho| \sigma_X \sigma_S \right) \right. \\ &\quad \left. - \gamma \left(\frac{\sigma_X^2 - r}{\kappa}, \frac{1}{\kappa} |\rho| \sigma_X \sigma_S e^{-\kappa T} \right) \right), \\ \Lambda_2(T) &= e^{(\sigma_X^2 - r)T} \left(\frac{|\rho| \sigma_X \sigma_S}{\kappa} \right)^{-\frac{\sigma_X^2 - r}{\kappa}} \left(\gamma \left(\frac{\sigma_X^2 - r}{\kappa} + 1, \frac{1}{\kappa} |\rho| \sigma_X \sigma_S \right) \right. \\ &\quad \left. - \gamma \left(\frac{\sigma_X^2 - r}{\kappa} + 1, \frac{1}{\kappa} |\rho| \sigma_X \sigma_S e^{-\kappa T} \right) \right), \\ \lambda(T) &= -S_0 e^{-\kappa T} - m(1 - e^{-\kappa T}) + \frac{\sigma_S^2}{4\kappa} (1 - e^{-2\kappa T}) - \frac{\rho \sigma_S \sigma_X}{\kappa} (1 - e^{-\kappa T}). \end{aligned}$$

Here $\gamma(s, x) = \int_0^x y^{s-1} e^{-y} dy$ denotes the incomplete Gamma function. Furthermore we have

$$\mathbb{V}(I_T) = A(2, 2, X_0, S_0, T) - A^2(1, 1, X_0, S_0, T),$$

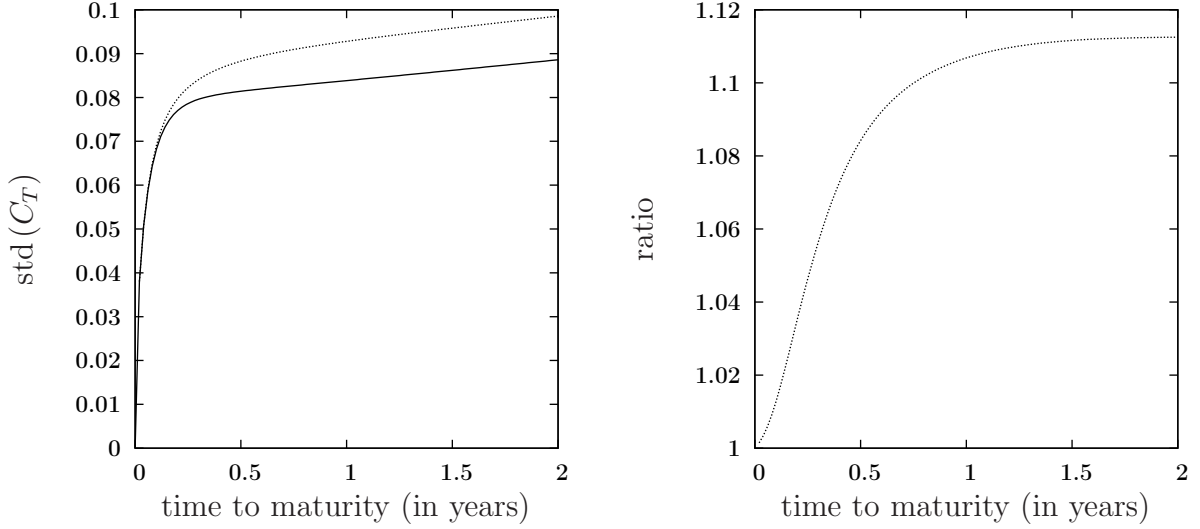


Figure 5: The panel on the left shows the standard deviation of the static (dashed line) versus the dynamic hedge error (solid line). The right hand panel depicts the ratio. As we have to assume a constant interest rate for the computation of the variance of the static hedge error we fix it at $r = 0.02$.

where A is as in (20).

Remark V.4. Note the assumption $\sigma_X^2 > r$ in the Proposition above is only needed in order to get a closed expression with respect to the incomplete Gamma function. In case $\sigma_X^2 \leq r$ the defining integral of the incomplete Gamma function explodes around 0. In this case the Formula (32) still holds if we replace γ with the *upper incomplete Gamma function* $\gamma(s, x) = -\int_x^\infty y^{s-1} e^{-y} dy$. In any case, regardless of $\sigma_X^2 > r$, the formula (32) holds when we replace $\Lambda_1(T) - \Lambda_2(T)$ with $\int_0^T e^{(\sigma_X^2 - r)t} (\sigma_X^2 - \rho\sigma_X\sigma_S e^{-\kappa(T-t)}) e^{-\rho\sigma_X\sigma_S e^{-\kappa(T-t)}/\kappa} dt$.

Proposition V.5. *The expressions for the optimal static hedge \check{a} (30) and for the corresponding variance (31) from the previous proposition hold also in the 2GBM case. For the involved*

expectations we have

$$\mathbb{E} \left(\int_0^T e^{-2rt} X_t^2 dt \right) = \begin{cases} \frac{X_0^2}{\sigma_X^2 - 2r} (e^{(\sigma_X^2 - 2r)T} - 1), & \text{if } \sigma_X^2 \neq 2r, \\ X_0^2 T, & \text{if } \sigma_X^2 = 2r, \end{cases}$$

and

$$\mathbb{E} \left(I_T \int_0^T e^{-rt} dX_t \right) = \begin{cases} X_0 I_0 \frac{\rho_{IX} \sigma_X \sigma_I}{\rho_{IX} \sigma_X \sigma_I - r} (e^{(\rho_{IX} \sigma_X \sigma_I - r)T} - 1), & \text{if } \rho_{IX} \sigma_X \sigma_I \neq r, \\ X_0 I_0 \rho_{IX} \sigma_X \sigma_I T, & \text{if } \rho_{IX} \sigma_X \sigma_I = r. \end{cases}$$

Furthermore

$$\mathbb{V}(I_T) = I_0^2 (e^{\sigma_I^2 T} - 1).$$

Using the expressions for the standard deviation of the hedge error from Theorem IV.2 and Proposition V.3 we can compare the risks entailed by both strategies. The left hand panel of Figure 5 depicts the standard deviation of the static and dynamic variance minimizing hedge against time to maturity. The right hand panel shows the increase of the standard deviation if one confines with the static hedge. The increase in the variability by more than 10% for positions hedged over a period of one year indicates that the hedge should be dynamically adjusted. The figure is again based on the estimated parameter values obtained for the August 2009 crude oil futures contract (see Table A1 in Appendix A).

VI. Including Directional Views and Stochastic Volatility

In this section we extend the model introduced in Section II by allowing for stochastic volatility of the futures and the logspread. We assume that the volatility of both processes are proportional to a Cox-Ingersoll-Ross process. The futures dynamics thus coincides with the dynamics

of the risky asset in the Heston model.

Let (W^1, W^2, W^3) be a 3-dimensional Brownian motion and suppose that the futures price process X and its volatility $\nu = (\nu_t)_{t \geq 0}$ evolve according to the SDE

$$(33) \quad dX_t = \mu(t, \nu_t)X_t dt + \sqrt{\nu_t}X_t dW_t^1$$

$$(34) \quad d\nu_t = \beta(\vartheta - \nu_t) dt + \sigma_\nu \sqrt{\nu_t}(\rho_1 dW_t^1 + \bar{\rho}_1 dW_t^2),$$

where $\rho_1 \in [-1, 1]$, $\bar{\rho}_1 = \sqrt{1 - \rho_1^2}$, $\beta, \vartheta, \sigma_\nu > 0$, and $\mu : \mathbb{R}_+^2 \rightarrow \mathbb{R}$ is measurable. As before, let $S_t = \log(X_t) - \log(I_t)$. Assume that the logspread's volatility is proportional to ν , and that S solves the mean reverting SDE

$$(35) \quad dS_t = \kappa(m - S_t) dt + \sigma_S \sqrt{\nu_t}(\rho dW_t^1 + \bar{\rho}\eta dW_t^2 + \bar{\rho}\bar{\eta} dW_t^3), \quad S_0 = s.$$

Since we have included a directional view in the dynamics of the futures price, we cannot directly invoke the method used in Section III for the derivation of the variance optimal hedge. When the trading instruments are assumed to be trended, and hence are not martingales, then it is very difficult to determine variance optimal hedging strategy. There are, however, other quadratic optimality criteria that considerably simplify the calculation of hedging strategies. A very intriguing type of hedging strategies are the so-called *locally risk minimizing hedging strategies*. These are variance optimal strategies with respect to a particular martingale measure, usually referred to as the *minimal martingale measure*. For an overview on quadratic hedging approaches we refer to Schweizer (2008).

In our extended model the minimal martingale measure \widehat{Q} is given by

$$\frac{d\widehat{Q}}{dP} = \exp\left(-\int_0^T \omega(t, \nu) dW_t^1 - \frac{1}{2} \int_0^T \omega^2(t, \nu) dt\right),$$

where $\omega(t, \nu_t) = \mu(t, \nu_t)/\sqrt{\nu_t}$ is the market price of risk.⁴

One can proceed as in Section III for the derivation of the local risk minimizing hedge, i.e. the variance optimal hedge relative to \widehat{Q} . The *value function* of $h(I_T)$ will be defined by

$$(36) \quad \psi(t, x, v, s) = e^{-r(T-t)} \mathbb{E}^{\widehat{Q}} \left(h(X_T^{t,x,v} e^{-S_T^{t,s}}) \right).$$

With the same assumptions on h one can show that the local risk minimizing hedge is given by

$$(37) \quad \widehat{\xi}_t = \psi_x(t, X_t, \nu_t, S_t) [1 - \sigma_S \rho e^{-\kappa(T-t)}] + \frac{\sigma_\nu \rho_1}{X_t} \psi_v(t, X_t, \nu_t, S_t).$$

Observe that the local risk minimizing hedge is now a weighted sum of the Delta and Vega of the risk position's expectation under the minimal martingale measure \widehat{Q} . Clearly, the term involving the Vega of the position appears due to the additional non-tradable risk induced by the stochastic volatility, which also needs to be cross-hedged.

A similar analysis as in the previous sections, e.g. estimation of model parameters, derivation of hedge errors and their respective standard deviations, is somewhat more involved. However, one can profit of the affine model structure and express the value function and its gradient in terms of generalized Ricatti equations. Fourier inversion methods then yield semi-explicit formulas for optimal strategies, which are amenable to swift simulation analysis.

VII. Conclusion and Outlook

Hedging is essential for controlling and managing risk and it is an important area of research. In this paper we show that a long-term relationship between the risk and the hedging instrument

⁴A sufficient condition for this to be a proper measure change is the following growth condition on ω . For

$A, B \geq 0$ and $\delta \in [0, 1/2]$ we assume $|\omega(t, x)| \leq A + Bx^\delta$, $x \geq 0$.

has important implications for the optimal hedging strategy and, thus, also for the hedge error. In particular, we propose a model which explicitly takes into account such a long-term relationship. We derive the variance optimal cross-hedge strategy and provide the variance of the hedge error in terms of the model's parameters. We demonstrate the practical relevance of incorporating the long-term relationship through an empirical example, where we find a long-term relationship between most crude oil futures contracts and the spot kerosene price. Interestingly, the model is also consistent with the commonly observed behavior of commodity traders, who use for cross hedges a hedge ratio of 100% instead of a hedge ratio dampened by the cross correlation between the risk and the hedging instrument, which is implied by models with only correlated Brownian motions. Furthermore, we show that even for cases where the decision concerning the stationarity of the logspread is not obvious, it is better to allow for a long-term relationship rather than to neglect it.

The model can be extended towards several directions to provide a more realistic dynamics for asset prices. Especially the consideration of jumps in the price process seems to be an interesting extension. However, several specifications are plausible and a careful empirical investigation is needed. On an ad hoc basis it is, for example, not clear whether the price processes jump together and how the jump sizes are related. These aspects will have significant impact on the properties of the hedge error and we plan to investigate these questions in more detail in future research.

Appendix A. Empirical Results

Table A1 shows the number of observations (second column), the p -value of the augmented Dickey-Fuller test (third column) and the estimation result for different crude oil futures contracts and their corresponding spot kerosene prices.

Table A1: Estimates

contract	nobs	ADF	μ	σ_x	σ_s	κ	m	ρ
201010	1147	0.0888*	0.0751 (0.1302)	0.2768 (0.0059)	0.2871 (0.0060)	2.5548 (1.0976)	-0.1661 (0.0534)	0.3401 (0.0260)
201009	1147	0.0750*	0.0742 (0.1324)	0.2795 (0.0059)	0.2877 (0.0060)	2.7131 (1.1078)	-0.1687 (0.0515)	0.3475 (0.0260)
201008	1145	0.0631*	0.0763 (0.1361)	0.2823 (0.0059)	0.2886 (0.0061)	2.8890 (1.1728)	-0.1693 (0.0492)	0.3550 (0.0258)
201007	1123	0.0555*	0.0778 (0.1348)	0.2855 (0.0061)	0.2908 (0.0062)	3.0632 (1.1704)	-0.1731 (0.0473)	0.3646 (0.0259)
201006	1375	0.0310**	0.1703 (0.1209)	0.2758 (0.0053)	0.2993 (0.0058)	2.5899 (1.1150)	-0.1946 (0.0493)	0.3404 (0.0241)
201005	1080	0.0436**	0.1114 (0.1405)	0.2901 (0.0064)	0.2942 (0.0064)	3.5046 (1.2913)	-0.1752 (0.0410)	0.3771 (0.0262)
201004	1058	0.0326**	0.0906 (0.1453)	0.2954 (0.0065)	0.2970 (0.0065)	3.8040 (1.3324)	-0.1808 (0.0390)	0.3867 (0.0262)
201003	1035	0.0250**	0.0737 (0.1494)	0.3002 (0.0066)	0.3006 (0.0067)	4.1174 (1.4241)	-0.1836 (0.0371)	0.3990 (0.0262)
201002	1015	0.0193**	0.0906 (0.1518)	0.3035 (0.0068)	0.3020 (0.0069)	4.5175 (1.6076)	-0.1866 (0.0332)	0.4043 (0.0263)
201001	994	0.0126**	0.0789 (0.1551)	0.3098 (0.0070)	0.3045 (0.0069)	4.9939 (1.3783)	-0.1916 (0.0312)	0.4178 (0.0262)
200912	1245	0.0097***	0.1852 (0.1345)	0.2979 (0.0060)	0.3114 (0.0064)	3.8232 (1.2755)	-0.2137 (0.0374)	0.3892 (0.0241)
200911	950	0.0046***	0.0860 (0.1564)	0.3197 (0.0073)	0.3108 (0.0072)	6.3841 (2.0112)	-0.1994 (0.0238)	0.4415 (0.0261)
200910	928	0.0029***	0.0603 (0.1687)	0.3229 (0.0075)	0.3135 (0.0074)	7.0739 (1.9253)	-0.2032 (0.0231)	0.4532 (0.0261)
200909	906	0.0014***	0.0819 (0.1738)	0.3275 (0.0077)	0.3178 (0.0076)	8.0814 (2.1057)	-0.2061 (0.0208)	0.4677 (0.0260)
200908	885	≤ 0.0010 ***	0.0430 (0.1807)	0.3321 (0.0079)	0.3223 (0.0078)	9.5437 (2.2822)	-0.2120 (0.0178)	0.4806 (0.0259)
200907	862	≤ 0.0010 ***	0.0743 (0.1828)	0.3388 (0.0082)	0.3275 (0.0080)	11.5957 (2.5298)	-0.2166 (0.0153)	0.5011 (0.0256)
200906	1114	≤ 0.0010 ***	0.1416 (0.1498)	0.3191 (0.0068)	0.3298 (0.0072)	6.6833 (1.7074)	-0.2404 (0.0233)	0.4579 (0.0239)
200905	819	≤ 0.0010 ***	-0.0114 (0.1953)	0.3516 (0.0086)	0.3407 (0.0086)	15.7614 (2.9919)	-0.2274 (0.0111)	0.5364 (0.0250)
200904	797	≤ 0.0010 ***	-0.0655 (0.1973)	0.3513 (0.0088)	0.3429 (0.0087)	15.9689 (3.0523)	-0.2307 (0.0103)	0.5594 (0.0245)
200903	775	≤ 0.0010 ***	-0.0672 (0.1953)	0.3436 (0.0086)	0.3351 (0.0086)	15.0723 (2.8192)	-0.2362 (0.0135)	0.5631 (0.0244)
200902	755	≤ 0.0010 ***	-0.0694 (0.2005)	0.3469 (0.0089)	0.3257 (0.0085)	12.1794 (2.4786)	-0.2411 (0.0146)	0.5815 (0.0242)
200901	733	0.0021***	-0.0802 (0.1966)	0.3306 (0.0086)	0.3053 (0.0081)	10.9528 (2.4631)	-0.2398 (0.0160)	0.5742 (0.0247)

The first column presents the maturity date of the contract, the second column gives the number of observations. The third column reports the p -value of the augmented Dickey-Fuller test for the null of non-stationarity. The * (**, ***) indicates the rejection of non-stationary at the 10% (5%, 1%) level. The remaining columns show the parameter estimates with the corresponding asymptotic standard errors given in parenthesis.

Appendix B. Proofs

Proof of Lemma II.1. By distinguishing the cases $\rho \geq 0$ and $\rho < 0$ one can show that $\partial\rho_{IX}/\partial\sigma_S \leq 0$, and that the partial derivative is strictly smaller than zero if $\sigma_S > 0$. Thus, ρ_{IX} is strictly decreasing in σ_S . From the definition of ρ_{IX} we have $\rho_{IX}^2 = (\sigma_X - \rho\sigma_S)^2 / (\sigma_X^2 - 2\rho\sigma_S\sigma_X + \sigma_S^2)$, which leads to the quadratic equation in σ_S

$$(B-1) \quad (\rho_{IX}^2 - \rho^2)\sigma_S^2 + 2\rho\sigma_X(1 - \rho_{IX}^2)\sigma_S - \sigma_X^2(1 - \rho_{IX}^2) = 0.$$

If $\rho \neq \rho_{IX}$, then Equation (B-1) has two solutions, namely

$$\sigma_S = \sigma_X \sqrt{1 - \rho_{IX}^2} \frac{-\rho\sqrt{1 - \rho_{IX}^2} \pm \rho_{IX}\sqrt{1 - \rho^2}}{\rho_{IX}^2 - \rho^2}.$$

Since $\rho_{IX}^2 - \rho^2 = (\rho_{IX}\sqrt{1 - \rho^2} + \rho\sqrt{1 - \rho_{IX}^2})(\rho_{IX}\sqrt{1 - \rho^2} - \rho\sqrt{1 - \rho_{IX}^2})$, this further yields

$$\sigma_S = \sigma_X \sqrt{1 - \rho_{IX}^2} \frac{1}{\rho\sqrt{1 - \rho_{IX}^2} \pm \rho_{IX}\sqrt{1 - \rho^2}}.$$

If $\rho_{IX} > \rho$, then only one of the roots guarantees that $\sigma_S \geq 0$, and we obtain (5). If $\rho_{IX} = \rho$, then (B-1) has a unique solution, given by $\sigma_S = \sigma_X/(2\rho)$. The inverse function of (4) is continuous on $\rho^{-1}(\mathbb{R}_+)$. Therefore, Equation (5) must also hold true for $\rho_{IX} < \rho$. \square

Proof of Lemma IV.1. Since $S_t^{0,s} = se^{-\kappa t} + m(1 - e^{-\kappa t}) + \int_0^t e^{-\kappa(t-u)}\sigma_S(\rho dW_u + \bar{\rho} dW_u^\perp)$, we get

$$\begin{aligned} e^{-aS_t^{0,s}}(X_t^{0,x})^b &= x^b \exp\left(-\frac{b}{2}\sigma_X^2 t - ase^{-\kappa t} - am(1 - e^{-\kappa t}) + \int_0^t (b\sigma_X - \rho a\sigma_S e^{-\kappa(t-u)}) dW_u\right) \\ &\quad \times \exp\left(-\int_0^t (a\bar{\rho}\sigma_S e^{-\kappa(t-u)}) dW_u^\perp\right). \end{aligned}$$

We calculate the variances of the independent normal variables given by the integrals in the

last two factors. We have

$$\int_0^t (b\sigma_X - \rho a\sigma_S e^{-\kappa(t-u)})^2 dt = b^2\sigma_X^2 t - 2ab\rho\sigma_X\sigma_S \frac{1}{\kappa}(1 - e^{-\kappa t}) + a^2\rho^2\sigma_S^2 \frac{1}{2\kappa}(1 - e^{-2\kappa t})$$

and

$$\int_0^t (a\bar{\rho}\sigma_S e^{-\kappa(t-u)})^2 dt = a^2\bar{\rho}^2\sigma_S^2 \frac{1}{2\kappa}(1 - e^{-2\kappa t}).$$

We use this to derive

$$\begin{aligned} \mathbb{E}\left(e^{-aS_t^{0,s}}(X_t^{0,x})^b\right) &= x^b \exp\left(-\frac{b}{2}\sigma_X^2 t - ase^{-\kappa t} - am(1 - e^{-\kappa t})\right) \\ &\quad \times \mathbb{E}\left(\exp\left(\int_0^t (b\sigma_X - \rho a\sigma_S e^{-\kappa(t-u)}) dW_u\right)\right) \\ &\quad \times \mathbb{E}\left(\exp\left(-\int_0^t (a\bar{\rho}\sigma_S e^{-\kappa(t-u)}) dW_u^\perp\right)\right) \\ &= x^b \exp\left(-\frac{b}{2}\sigma_X^2 t - ase^{-\kappa t} - am(1 - e^{-\kappa t})\right) \\ &\quad \times \exp\left[\frac{1}{2}\left(b^2\sigma_X^2 t - 2ab\rho\sigma_X\sigma_S \frac{1}{\kappa}(1 - e^{-\kappa t}) + a^2\rho^2\sigma_S^2 \frac{1}{2\kappa}(1 - e^{-2\kappa t})\right)\right] \\ &\quad \times \exp\left[\frac{1}{2}\left(a^2\bar{\rho}^2\sigma_S^2 \frac{1}{2\kappa}(1 - e^{-2\kappa t})\right)\right], \end{aligned}$$

from which the result follows. □

Proof of Theorem IV.2. Recall that from (21) we have

$$(B-2) \quad \mathbb{V}(C_T(\xi^*, v)) = c^2\bar{\rho}^2\sigma_S^2 \int_0^T e^{-2k(T-t)} \mathbb{E}(X_t^2 A^2(1, 1, 1, S_t, T-t)) dt.$$

By the definition of A , we have

(B-3)

$$\begin{aligned} \mathbb{E} \left(X_t^2 A^2(1, 1, 1, S_t, T-t) \right) &= \exp \left(-2(m + \rho \sigma_X \sigma_S \frac{1}{\kappa})(1 - e^{-\kappa(T-t)}) + \sigma_S^2 \frac{1}{2\kappa}(1 - e^{-2\kappa(T-t)}) \right) \\ &\quad \times \mathbb{E} \left(X_t^2 \exp(-2S_t e^{-\kappa(T-t)}) \right), \end{aligned}$$

and again using the definition of A we get

$$\begin{aligned} \mathbb{E} \left(X_t^2 \exp(-2S_t e^{-\kappa(T-t)}) \right) &= A(2e^{-\kappa(T-t)}, 2, x, s, t) \\ &= x^2 \exp \left[\sigma_X^2 t - 2s e^{-\kappa T} - 2(m + 2\rho \sigma_X \sigma_S \frac{1}{\kappa})(e^{-\kappa(T-t)} - e^{-\kappa T}) \right] \\ &\quad \times \exp \left[2\sigma_S^2 \frac{1}{2\kappa}(e^{-2\kappa(T-t)} - e^{-2\kappa T}) \right]. \end{aligned}$$

Combining the last equation with (B-3) we further obtain

(B-4)

$$\begin{aligned} \mathbb{E} \left(X_t^2 A^2(1, 1, 1, S_t, T-t) \right) &= x^2 \exp \left(\sigma_X^2 t - 2s e^{-\kappa T} - 2\rho \sigma_X \sigma_S \frac{1}{\kappa}(1 + e^{-\kappa(T-t)} - 2e^{-\kappa T}) \right) \\ &\quad \times \exp \left(-2m(1 - e^{-\kappa T}) + \sigma_S^2 \frac{1}{2\kappa}(1 - 2e^{-\kappa T} + e^{-2\kappa(T-t)}) \right). \end{aligned}$$

The previous calculations yield, by combination of (B-2) and (B-4), Equation (22). \square

Proof of Proposition V.1. Since I is a GBM in the 2GBM model, the value function ψ associated with the linear position $h(x) = cx$ is given by $\psi(t, y) = e^{-r(T-t)}cy$. Therefore, $\psi_y(t, y) = e^{-r(T-t)}c$, and the realized error variance in Model 1, following strategy $\zeta_t = \rho_{IX}\sigma_I I_t e^{-r(T-t)}c / (\sigma_X X_t)$, satisfies

$$C_T = cI_T - e^{rT} \left(v + \int_0^T e^{-rT} c \rho_{IX} \sigma_I I_t dW_t^{(X)} \right),$$

and consequently we have

$$\mathbb{V}(C_T) = c^2 \mathbb{V}(I_T) - 2c^2 \text{Cov} \left(I_T, \int_0^T \rho_{IX} \sigma_I I_t dW_t^{(X)} \right) + c^2 \mathbb{E} \left(\int_0^T \rho_{IX}^2 \sigma_I^2 I_t^2 dt \right).$$

Note that $\mathbb{V}(I_T) = A(2, 2, X_0, S_0, T) - A^2(1, 1, X_0, S_0, T)$ and observe further that

$$\mathbb{E} \left(\int_0^T \rho_{IX}^2 \sigma_I^2 I_t^2 dt \right) = \int_0^T \rho_{IX}^2 \sigma_I^2 A(2, 2, X_0, S_0, t) dt.$$

It remains to calculate the covariance above. To that effect we recall the decomposition (B-5) of I_T from the proof of Proposition V.3 and borrow the respective notation to write

$$\int_0^T \rho_{IX} \sigma_I I_t dW_t^{(X)} = \int_0^T \rho_{IX} \sigma_I I_t d\tilde{W}_t^{(X)} + \int_0^T \rho_{IX} \sigma_I I_t (\sigma_X - \rho \sigma_S e^{-\kappa(T-t)}) dt. \text{ Consequently,}$$

$$\begin{aligned} \text{Cov} \left(I_T, \int_0^T \rho_{IX} \sigma_I I_t dW_t^{(X)} \right) &= X_0 e^{\lambda(T)} \mathbb{E}^Q \left(\int_0^T \rho_{IX} \sigma_I I_t (\sigma_X - \rho \sigma_S e^{-\kappa(T-t)}) dt \right) \\ &= X_0 e^{\lambda(T)} \int_0^T \rho_{IX} \sigma_I (\sigma_X - \rho \sigma_S e^{-\kappa(T-t)}) \mathbb{E}^Q(I_t) dt \\ &= \int_0^T \rho_{IX} \sigma_I (\sigma_X - \rho \sigma_S e^{-\kappa(T-t)}) \mathbb{E}(I_T I_t) dt. \end{aligned}$$

In order to calculate $\mathbb{E}(I_T I_t)$ we proceed similar as in Lemma IV.1. We decompose $I_T I_t$ into

$$\begin{aligned} I_T I_t &= X_0^2 \exp \left(\int_0^T [\sigma_X (1 + 1_{u \leq t}) - \sigma_S \rho (e^{-\kappa(T-u)} + e^{-\kappa(t-u)} 1_{u \leq t})] dW_u^X \right) \\ &\quad \times \exp \left(- \int_0^T \sigma_S \bar{\rho} (e^{-\kappa(T-u)} + e^{-\kappa(t-u)} 1_{u \leq t}) dW_u^\perp \right) \\ &\quad \times \exp \left(- \frac{1}{2} \int_0^T \sigma_X^2 (1 + 1_{u \leq t}) du \right) \exp(-S_0 (e^{-\kappa T} + e^{-\kappa t}) - m(2 - e^{-\kappa T} - e^{-\kappa t})). \end{aligned}$$

The variances of the stochastic integrals are given by

$$\begin{aligned} &\int_0^T \sigma_X^2 (1 + 2_{u \leq t} + 1_{u \leq t}) - 2\sigma_S \sigma_X \rho [e^{-\kappa(T-u)} + 1_{u \leq t} (2e^{-\kappa(t-u)} + e^{-\kappa(T-u)})] \\ &\quad + \sigma_S^2 \rho^2 (e^{-\kappa(T-u)} + e^{-\kappa(t-u)} 1_{u \leq t})^2 du \end{aligned}$$

and

$$\int_0^T \sigma_S^2 \bar{\rho}^2 \left(e^{-\kappa(T-u)} + e^{-\kappa(t-u)} \mathbf{1}_{u \leq t} \right)^2 \mathrm{d}u.$$

Hence taking expectation yields

$$\begin{aligned} \mathbb{E}(I_T I_t) &= X_0^2 \exp \left(\frac{1}{2} \int_0^T \left[\sigma_X^2 2_{u \leq t} - 2\sigma_S \sigma_X \rho \left[e^{-\kappa(T-u)} + \mathbf{1}_{u \leq t} \left(2e^{-\kappa(t-u)} + e^{-\kappa(T-u)} \right) \right] \right] \mathrm{d}u \right) \\ &\quad \times \exp \left(\frac{1}{2} \int_0^T \sigma_S^2 \left(e^{-2\kappa(T-u)} + 2e^{-\kappa(T+t-2u)} \mathbf{1}_{u \leq t} + e^{-2\kappa(t-u)} \mathbf{1}_{u \leq t} \right) \mathrm{d}u \right) \\ &\quad \times \exp \left(-S_0 \left(e^{-\kappa T} + e^{-\kappa t} \right) - m \left(2 - e^{-\kappa T} - e^{-\kappa t} \right) \right). \end{aligned}$$

The result follows by a simple calculation. \square

Proof of Proposition V.2. Recall from (13) that $\xi_t^* = [1 - \sigma_S \rho e^{-\kappa(T-t)} / \sigma_X] \psi_x(t, X_t, S_t)$, with $\psi(t, x, s) = e^{-r(T-t)} \mathbb{E} \left(h(X_T^{t,x} e^{-S_T^{t,s}}) \right)$.

Hence for $h(y) = cy$ we get $\psi_x(t, x, s) = ce^{-r(T-t)} \mathbb{E} \left(X_T^{t,1} e^{-S_T^{t,s}} \right) = A(1, 1, 1, s, T-t)$. Thus, the realized error variance in the 2GBM model, following the strategy above, satisfies

$$C_T = cI_T - e^{rT} \left(v + c \int_0^T e^{-rT} \left[1 - \frac{\sigma_S}{\sigma_X} \rho e^{-\kappa(T-t)} \right] A(1, 1, 1, S_t, T-t) \mathrm{d}X_t \right).$$

Consequently, setting $v = ce^{-rT} \mathbb{E}(I_T)$, we have

$$\begin{aligned} \mathbb{V}(C_T) &= c^2 \mathbb{V}(I_T) - 2c^2 \mathrm{Cov} \left(I_T, \int_0^T \left[1 - \frac{\sigma_S}{\sigma_X} \rho e^{-\kappa(T-t)} \right] A(1, 1, 1, S_t, T-t) \sigma_X X_t \mathrm{d}W_t^{(X)} \right) \\ &\quad + c^2 \mathbb{E} \left(\int_0^T \left[1 - \frac{\sigma_S}{\sigma_X} \rho e^{-\kappa(T-t)} \right]^2 A^2(1, 1, 1, S_t, T-t) \sigma_X^2 X_t^2 \mathrm{d}t \right). \end{aligned}$$

It is straightforward to see that $B(a, b, t)$ defined as $B(a, b, t) = \mathbb{E} \left(X_t^a I_t^b \right)$, with X and I as in the 2GBM model, fulfills (29). Hence $\mathbb{V}(I_T) = B(0, 2, T) - B^2(0, 1, T)$. Observe that we may,

via Fubini's theorem, write the expectation in the last term in the variance of C_T above as

$$\int_0^T \left[1 - \frac{\sigma_S}{\sigma_X} \rho e^{-\kappa(T-t)} \right]^2 A^2(1, 1, 1, 0, T-t) \sigma_X^2 \mathbb{E} \left(\exp(-2S_t e^{-\kappa(T-t)}) X_t^2 \right) dt.$$

In order to simplify further recall that $S_t = \log(X_t) - \log(I_t)$. Thus,

$$\mathbb{E} \left(\exp(-2S_t e^{-\kappa(T-t)}) X_t^2 \right) = \mathbb{E} \left(X_t^{2-2e^{-\kappa(T-t)}} I_t^{2e^{-\kappa(T-t)}} \right) = B(2 - 2e^{-\kappa(T-t)}, 2e^{-\kappa(T-t)}, t),$$

which combined with the previous integral yields the last term in (28). In order to simplify the remaining term in the variance of C_T we apply the standard trick of a change of measure, here with $dQ = (I_T/I_0) dP$, Fubini's Theorem, and reversing the measure change to obtain

$$\int_0^T \sigma_I \rho \left[1 - \frac{\sigma_S}{\sigma_X} \rho e^{-\kappa(T-t)} \right] A(1, 1, 1, 0, T-t) \sigma_X \mathbb{E} \left(I_T \exp(-S_t e^{-\kappa(T-t)}) X_t \right) dt.$$

Finally, using $S_t = \log(X_t) - \log(I_t)$, and the independent increments of a Brownian motion, we get

$$\begin{aligned} \mathbb{E} \left(I_T \exp(-S_t e^{-\kappa(T-t)}) X_t \right) &= \mathbb{E} \left(I_T I_t^{e^{-\kappa(T-t)}} X_t^{1-e^{-\kappa(T-t)}} \right) \\ &= \mathbb{E} \left(I_t^{1+e^{-\kappa(T-t)}} X_t^{1-e^{-\kappa(T-t)}} \exp \left(\int_t^T \sigma_I dW_u^I - \frac{1}{2} \int_t^T \sigma_I^2 du \right) \right) \\ &= B(1 - e^{-\kappa(T-t)}, 1 + e^{-\kappa(T-t)}, t), \end{aligned}$$

which together with the previous integral yields the middle term in (28) and thus finishes the proof. \square

Proof of Proposition V.3. The variance of the hedge error does not depend on the initial capital v , and hence we may assume that $e^{rT}v = c\mathbb{E}(I_T)$. Holding the constant position a

between 0 and T then entails the hedge error

$$C_T(a) = cI_T - \mathbb{E}(cI_T) - ae^{rT} \int_0^T e^{-rt} dX_t,$$

and since X_t is a martingale we have $\mathbb{E}(C_T(a)) = 0$. Hence, the variance of $C_T(a)$ is given by

$$\begin{aligned} \mathbb{E}(C_T^2(a)) &= c^2 \mathbb{E}((I_T - \mathbb{E}(I_T))^2) - 2ace^{rT} \mathbb{E}\left((I_T - \mathbb{E}(I_T)) \int_0^T e^{-rt} dX_t\right) \\ &\quad + a^2 e^{2rT} \mathbb{E}\left(\left(\int_0^T e^{-rt} dX_t\right)^2\right) \\ &= c^2 \mathbb{E}((I_T - \mathbb{E}(I_T))^2) - 2ace^{rT} \mathbb{E}\left(I_T \int_0^T e^{-rt} dX_t\right) + a^2 e^{2rT} \mathbb{E}\left(\int_0^T e^{-2rt} \sigma_X^2 X_t^2 dt\right). \end{aligned}$$

The optimal \check{a} which minimizes the variance of the hedge error is given by

$$\check{a} = ce^{-rT} \frac{\mathbb{E}\left(I_T \int_0^T e^{-rt} dX_t\right)}{\sigma_X^2 \mathbb{E}\left(\int_0^T e^{-2rt} X_t^2 dt\right)}.$$

For the variance of the corresponding hedge error we have

$$\mathbb{E}(C_T^2(\check{a})) = \mathbb{V}(cI_T) - \frac{\text{Cov}\left(cI_T, e^{rT} \int_0^T e^{-rt} dX_t\right)^2}{\mathbb{V}\left(e^{rT} \int_0^T e^{-rt} dX_t\right)} = c^2 \mathbb{V}(I_T) - \frac{c^2 \left[\mathbb{E}\left(I_T \int_0^T e^{-rt} dX_t\right)\right]^2}{\sigma_X^2 \mathbb{E}\left(\int_0^T e^{-2rt} X_t^2 dt\right)}.$$

The expectation in the denominator is given by

$$\mathbb{E}\left(\int_0^T e^{-2rt} X_t^2 dt\right) = \int_0^T e^{-2rt} \underbrace{\mathbb{E}(X_t^2)}_{=X_0^2 \exp(\sigma_X^2 t)} dt = \begin{cases} \frac{X_0^2}{\sigma_X^2 - 2r} (e^{(\sigma_X^2 - 2r)T} - 1), & \text{if } \sigma_X^2 \neq 2r, \\ X_0^2 T, & \text{if } \sigma_X^2 = 2r. \end{cases}$$

The computations for the expectation in the numerator are somewhat more involved. Using

the explicit expressions for S_T and X_T , we may decompose I_T into

$$(B-5) \quad I_T = X_T e^{-S_T} = X_0 e^{\lambda(T)} D_T,$$

where D_T is the value at time T of the process D defined by, for all $t \in [0, T]$,

$$\begin{aligned} D_t &= \exp \left(\int_0^t (\sigma_X - \rho \sigma_S e^{-\kappa(T-u)}) dW_u^{(X)} - \frac{1}{2} \int_0^t (\sigma_X - \rho \sigma_S e^{-\kappa(T-u)})^2 du \right) \\ &\quad \times \exp \left(- \int_0^t \bar{\rho} \sigma_S e^{-\kappa(T-u)} dW_u^\perp - \frac{1}{2} \int_0^t \bar{\rho}^2 \sigma_S^2 e^{-2\kappa(T-u)} du \right) \end{aligned}$$

and $\lambda(T)$ is a constant given by

$$\begin{aligned} \lambda(T) &= -S_0 e^{-\kappa T} - m(1 - e^{-\kappa T}) - \frac{\sigma_X^2}{2} T + \frac{1}{2} \int_0^T (\sigma_X - \rho \sigma_S e^{-\kappa(T-u)})^2 du \\ &\quad + \frac{1}{2} \int_0^T \bar{\rho}^2 \sigma_S^2 e^{-2\kappa(T-u)} du \\ &= -S_0 e^{-\kappa T} - m(1 - e^{-\kappa T}) + \frac{\sigma_S^2}{4\kappa} (1 - e^{-2\kappa T}) - \frac{\rho \sigma_S \sigma_X}{\kappa} (1 - e^{-\kappa T}). \end{aligned}$$

Note that D is a strictly positive martingale and satisfies Novikov's condition. Therefore we can define a probability measure Q via

$$dQ = D_T dP.$$

Under Q the processes $\tilde{W}^{(X)}$ and \tilde{W}^\perp , for all $t \in [0, T]$,

$$\begin{aligned} \tilde{W}_t^{(X)} &= W_t^{(X)} - \int_0^t (\sigma_X - \rho \sigma_S e^{-\kappa(T-u)}) du \\ \tilde{W}_t^\perp &= W_t^\perp + \int_0^t \bar{\rho} \sigma_S e^{-\kappa(T-u)} du \end{aligned}$$

are independent Brownian motions. The dynamics of X , rewritten in terms of $\tilde{W}^{(X)}$, satisfy

$$dX_t = \sigma_X(\sigma_X - \rho\sigma_S e^{-\kappa(T-t)})X_t dt + \sigma_X X_t d\tilde{W}_t^{(X)}.$$

Observe that the expectation of X_t with respect to Q is given by

$$\begin{aligned} \mathbb{E}^Q(X_t) &= X_0 \exp\left(\int_0^t \sigma_X(\sigma_X - \rho\sigma_S e^{-\kappa(T-u)}) du\right) \\ &= X_0 \exp\left(\sigma_X^2 t + \frac{\rho\sigma_X\sigma_S}{\kappa} e^{-\kappa T}\right) \exp\left(-\frac{\rho\sigma_X\sigma_S}{\kappa} e^{-\kappa(T-t)}\right). \end{aligned}$$

Now the expectation term in the numerator can be written as

$$\begin{aligned} \mathbb{E}\left(I_T \int_0^T e^{-rt} dX_t\right) &= X_0 e^{\lambda(T)} \mathbb{E}^Q\left(\int_0^T e^{-rt} dX_t\right) \\ &= X_0 e^{\lambda(T)} \int_0^T e^{-rt} \sigma_X(\sigma_X - \rho\sigma_S e^{-\kappa(T-t)}) \mathbb{E}^Q(X_t) dt \\ \text{(B-6)} \quad &= X_0^2 e^{\lambda(T) + \frac{\rho\sigma_X\sigma_S}{\kappa} e^{-\kappa T}} \\ &\quad \times \underbrace{\int_0^T e^{(\sigma_X^2 - r)t} (\sigma_X^2 - \rho\sigma_X\sigma_S e^{-\kappa(T-t)}) e^{-\frac{\rho\sigma_X\sigma_S}{\kappa} e^{-\kappa(T-t)}} dt}_{=A}. \end{aligned}$$

For $\rho = 0$ we are done. For $\rho \neq 0$ we continue by substituting $u = |\rho|\sigma_X\sigma_S e^{-\kappa(T-t)}/\kappa$ which leads to an explicit expression for the above integral A in terms of the incomplete Gamma

function $\gamma(s, x) = \int_0^x y^{s-1} e^{-y} dy$.

$$\begin{aligned}
A &= e^{(\sigma_X^2 - r)T} \left(\frac{|\rho| \sigma_X \sigma_S}{\kappa} \right)^{-\frac{\sigma_X^2 - r}{\kappa}} \int_{\frac{1}{\kappa} |\rho| \sigma_X \sigma_S e^{-\kappa T}}^{\frac{1}{\kappa} |\rho| \sigma_X \sigma_S} u^{\frac{\sigma_X^2 - r}{\kappa}} (\sigma_X^2 - \kappa u) e^{-u} \frac{1}{\kappa u} du \\
&= \frac{\sigma_X^2 e^{(\sigma_X^2 - r)T} (|\rho| \sigma_X \sigma_S)^{-\frac{\sigma_X^2 - r}{\kappa}}}{\kappa^{1 - \frac{\sigma_X^2 - r}{\kappa}}} \int_{\frac{1}{\kappa} |\rho| \sigma_X \sigma_S e^{-\kappa T}}^{\frac{1}{\kappa} |\rho| \sigma_X \sigma_S} u^{\frac{\sigma_X^2 - r}{\kappa} - 1} e^{-u} du \\
&\quad - e^{(\sigma_X^2 - r)T} \left(\frac{|\rho| \sigma_X \sigma_S}{\kappa} \right)^{-\frac{\sigma_X^2 - r}{\kappa}} \int_{\frac{1}{\kappa} |\rho| \sigma_X \sigma_S e^{-\kappa T}}^{\frac{1}{\kappa} |\rho| \sigma_X \sigma_S} u^{\frac{\sigma_X^2 - r}{\kappa}} e^{-u} du \\
&= \underbrace{\frac{\sigma_X^2 e^{(\sigma_X^2 - r)T} (|\rho| \sigma_X \sigma_S)^{-\frac{\sigma_X^2 - r}{\kappa}}}{\kappa^{1 - \frac{\sigma_X^2 - r}{\kappa}}} \left(\gamma \left(\frac{\sigma_X^2 - r}{\kappa}, \frac{1}{\kappa} |\rho| \sigma_X \sigma_S \right) - \gamma \left(\frac{\sigma_X^2 - r}{\kappa}, \frac{1}{\kappa} |\rho| \sigma_X \sigma_S e^{-\kappa T} \right) \right)}_{=\Lambda_1(T)} \\
&\quad - \underbrace{e^{(\sigma_X^2 - r)T} \left(\frac{|\rho| \sigma_X \sigma_S}{\kappa} \right)^{-\frac{\sigma_X^2 - r}{\kappa}} \left(\gamma \left(\frac{\sigma_X^2 - r}{\kappa} + 1, \frac{|\rho| \sigma_X \sigma_S}{\kappa} \right) - \gamma \left(\frac{\sigma_X^2 - r}{\kappa} + 1, \frac{|\rho| \sigma_X \sigma_S e^{-\kappa T}}{\kappa} \right) \right)}_{=\Lambda_2(T)}.
\end{aligned}$$

Plugging this into (B-6) gives

$$\mathbb{E} \left(I_T \int_0^T e^{-rt} dX_t \right) = X_0^2 e^{\lambda(T) + \frac{\rho \sigma_X \sigma_S}{\kappa} e^{-\kappa T}} (\Lambda_1(T) - \Lambda_2(T)).$$

The expression for the $\mathbb{V}(I_T)$ is straightforward. \square

Proof of Proposition V.5. The expressions for \check{a} and the corresponding variance $\mathbb{V}(C_T(\check{a}))$ are model independent and therefore the same as in Proposition V.3. In both models X is a GBM and hence we again have

$$\mathbb{E} \left(\int_0^T e^{-2rt} X_t^2 dt \right) = \begin{cases} \frac{X_0^2}{\sigma_X^2 - 2r} (e^{(\sigma_X^2 - 2r)T} - 1) & \text{if } \sigma_X^2 \neq 2r \\ X_0^2 T & \text{if } \sigma_X^2 = 2r \end{cases}.$$

For the expectation in the numerator we get

$$\mathbb{E} \left(I_T \int_0^T e^{-rt} dX_t \right) = \mathbb{E} \left(\int_0^T e^{-rt} d\langle I, X \rangle_t \right) = \int_0^T e^{-rt} \rho_{IX} \sigma_X \sigma_I \mathbb{E}(I_t X_t) dt.$$

Straightforward computations give $\mathbb{E}(X_t I_t) = X_0 I_0 e^{\rho_{IX} \sigma_X \sigma_I t}$, which plugged into the above expressions gives the desired result. The expression for the $\mathbb{V}(I_T)$ is straightforward. \square

References

- Alexander, C. “Optimal Hedging Using Cointegration.” *Philosophical Transactions of the Royal Society Series A*, 357 (1999), 2039–2058.
- Ankirchner, S., and G. Heyne. “Cross Hedging with Stochastic Correlation.” *Finance and Stochastics*, forthcoming (2011).
- Ankirchner, S.; P. Imkeller; and G. Dos Reis. “Pricing and Hedging of Derivatives Based on Nontradable Underlyings.” *Mathematical Finance*, 20 (2010), 289–312.
- Baillie, R. T. “Econometric Tests of Rationality and Market Efficiency.” *Econometric Reviews*, 8 (1989), 151–186.
- Baillie, R. T., and R. J. Myers. “Bivariate GARCH Estimation of the Optimal Commodity Futures Hedge.” *Journal of Applied Econometrics*, 6 (1991), 109–124.
- Brenner, R. J., and K. F. Kroner. “Arbitrage, Cointegration, and Testing the Unbiasedness Hypothesis in Financial Markets.” *Journal of Financial and Quantitative Analysis*, 30 (1995), 23–42.
- Casassus, J., and P. Collin-Dufresne. “Stochastic Convenience Yield Implied from Commodity Futures and Interest Rates.” *Journal of Finance*, 60 (2005), 2283–2331.

- Cecchetti, S. G.; R. E. Cumby; and S. Figlewski. “Estimation of the Optimal Futures Hedge.” *The Review of Economics and Statistics*, 70 (1988), 623–630.
- Chan, K. “A Further Analysis of the Lead-Lag Relationship Between the Cash Market and Stock Index Futures Market.” *Review of Financial Studies*, 5 (1987), 123–152.
- Davis, M. H. A. “Optimal Hedging with Basis Risk.” In *From Stochastic Calculus to Mathematical Finance*, Y. Kabanov; R. Liptser; and J. Stoyanov, eds. Berlin: Springer (2006).
- Duffie, D., and H. Richardson. “Mean-Variance Hedging in Continuous Time.” *Annals of Applied Probability*, 1 (1991), 1–15.
- Engle, R. F., and C. W. J. Granger. “Co-Integration and Error Correction: Representation, Estimation and Testing.” *Econometrica*, 55 (1987), 251–276.
- Föllmer, H., and D. Sondermann. “Hedging of Nonredundant Contingent Claims.” In *Contributions to Mathematical Economics*, A. Mas-Colell, and W. Hildenbrand, eds. Amsterdam: North-Holland (1986).
- Granger, C. W. J. “Some Properties of Time Series Data and Their Use in Econometric Model Specification.” *Journal of Econometrics*, 16 (1981), 121–130.
- Howard, C. T., and L. J. D’Antonio. “Multiperiod Hedging Using Futures: A Risk Minimization Approach in the Presence of Autocorrelation.” *Journal of Futures Markets*, 11 (1991), 697–710.
- Hull, J. C. *Options, Futures, and Other Derivatives*. Upper Saddle River: Prentice Hall, Seventh edition (2008).
- Hulley, H., and T. A. McWalter. “Quadratic Hedging of Basis Risk.” Research Paper Series 225, Quantitative Finance Research Centre, University of Technology, Sydney (2008).

- Karatzas, I., and S. E. Shreve. *Brownian Motion and Stochastic Calculus*. Berlin: Springer, second edition (1991).
- Kawaller, I. G.; P. D. Koch; and T. W. Koch. “The Temporal Price Relationship Between S&P 500 Futures and the S&P 500 Index.” *Journal of Finance*, 42 (1987), 1309–1329.
- Lien, D. “Optimal Hedging and Spreading in Cointegrated Markets.” *Economics Letters*, 40 (1992), 91–95.
- Lien, D. “Cointegration and the Optimal Hedge Ratio: The General Case.” *The Quarterly Review of Economics and Finance*, 44 (2004), 654–658.
- Lien, D., and X. Luo. “Estimating Multiperiod Hedge Ratios in Cointegrated Markets.” *Journal of Futures Markets*, 13 (1993), 909–920.
- Monoyios, M. “Performance of Utility-Based Strategies for Hedging Basis Risk.” *Quantitative Finance*, 4 (2004), 245–255.
- Musiela, M., and T. Zariphopoulou. “An Example of Indifference Prices Under Exponential Preferences.” *Finance and Stochastics*, 8 (2004), 229–239.
- Ng, V. K., and S. C. Pirrong. “Price Dynamics in Refined Petroleum Spot and Futures Markets.” *Journal of Empirical Finance*, 2 (1996), 359–388.
- Schwartz, E. S. “The Stochastic Behavior of Commodity Prices: Implications for Valuation and Hedging.” *Journal of Finance*, 52 (1997), 923–73.
- Schweizer, M. “Mean-Variance Hedging for General Claims.” *Annals of Applied Probability*, 2 (1992), 171–179.
- Schweizer, M. “Local Risk-Minimization for Multidimensional Assets and Payment Streams.” In *Advances in Mathematics of Finance*, Łukasz Stettner, ed. Warsaw: Banach Center Publications (2008).

Stoll, H. R., and R. E. Whaley. “The Dynamics of Stock Index and Stock Index Futures Returns.” *Journal of Financial and Quantitative Analysis*, 25 (1990), 441–468.