

# Pricing of long-dated commodity derivatives: Do stochastic interest rates matter?

Benjamin Cheng<sup>a</sup>, Christina Sklibosios Nikitopoulos<sup>a</sup>, Erik Schlögl<sup>a</sup>

<sup>a</sup>*University of Technology Sydney,  
Finance Discipline Group, UTS Business School,  
PO Box 123 Broadway NSW 2007, Australia*

---

## Abstract

Does modelling stochastic interest rates, beyond stochastic volatility, improve pricing performance on long-dated commodity derivatives? To answer this question, we consider futures price models for commodity derivatives that allow for stochastic volatility and stochastic interest rates and a correlation structure between the underlying variables. We examine the empirical pricing performance of these models on pricing long-dated crude oil derivatives. Estimating the model parameters from historical crude oil futures prices and option prices, we find that stochastic interest rate models improve pricing performance on long-dated crude oil derivatives, when the interest rate volatility is relatively high. Furthermore, increasing the model dimensionality does not tend to improve the pricing performance on long-dated crude oil option prices, but it matters for long-dated futures prices. We also find empirical evidence for a negative correlation between crude oil futures prices and interest rates that contributes to improving fit to long-dated crude oil option prices.

*Keywords:* Futures options pricing; Stochastic interest rates; Correlations; Long-dated crude oil derivatives;

*JEL:* C13, C60, G13, Q40

---

---

*Email addresses:* Benjamin.T.Cheng@student.uts.edu.au (Benjamin Cheng), christina.nikitopoulos@uts.edu.au (Christina Sklibosios Nikitopoulos), erik.schlogl@uts.edu.au (Erik Schlögl)

## 1. Introduction

The role of commodity markets in the financial sector has increased in importance over the last decade. A record high of \$277 billion invested in commodity exchange-traded products was observed in 2009<sup>1</sup> (which was 50 times larger than the decade earlier) with the crude oil market being the most active commodity market. A variety of new products have become available, including exchange-traded products, managed futures funds, and hedge funds that boost activity in both short-term trading and long-term investment strategies. Commodities markets along with real estate are becoming promising alternative investment vehicles beyond equity markets, with commodity indices outperforming the S&P 500 index over the last decade.

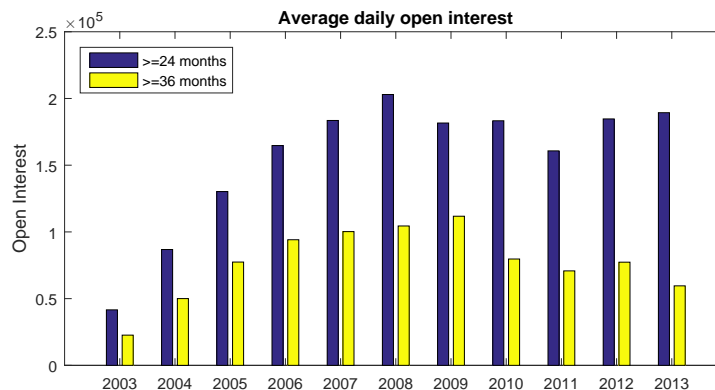


Figure 1: **Average daily open interest.** The average daily open interest of crude oil futures contracts with maturities more than or equal to 24 months (for  $\geq 24$ ) or 36 months (for  $\geq 36$ ) for each year from 2003 to 2013.

The crude oil futures and options are the world's most actively traded commodity derivatives, forming a major part of these activities. The average daily open interest in crude oil futures contracts of all maturities has increased from 503,662 contracts in 2003 to 1,677,627 contracts in 2013.<sup>2</sup> Even though the most active contracts are short-dated, the market for long-dated contracts has also substantially increased. The maturities of the crude oil futures contracts and the options on futures contracts have extended from 18 months in 1990 to over 9 years in recent years. The average daily open interest in crude oil futures contracts with

<sup>1</sup>Source: [www.barchart.com/articles/etf/commodityindex](http://www.barchart.com/articles/etf/commodityindex)

<sup>2</sup>The crude oil derivatives database, which includes information about the open interest and volume, was obtained from the Chicago Mercantile Exchange (CME).

maturities of two years or more was 41,601 contracts in December 2003, and reached a record high of 202,964 contracts in 2008. Figure 1 displays the average daily open interest<sup>3</sup> of crude oil futures contracts for each year from 2003 to 2013 and reveals a considerable trading activity for long-dated crude oil futures contracts in recent years. Motivated by the increasing importance of long-dated commodity derivative contracts to the financial markets, we make theoretical and empirical contributions on the pricing of long-dated commodity derivatives.

Studies dealing with pricing of long-dated derivatives typically use spot price models with stochastic interest rates and/or stochastic volatility and are applied mostly in equity or FX markets. Using European stock index options with maturities of up to two years from the Swedish option market, Rindell (1995) demonstrates empirically that the stochastic interest rate option pricing model of Amin and Jarrow (1992) outperforms the original Black and Scholes (1973) model. Some of the early models do not include correlations, or a sufficient number of factors, and many do not derive closed form option prices, including Amin and Jarrow (1992), Amin and Ng (1993), Bakshi, Cao, and Chen (1997), and Grzelak and Oosterlee (2011). By using long-term equity anticipation securities with maturities up to three years, Bakshi, Cao, and Chen (2000) empirically demonstrate that stochastic volatility and stochastic interest rate models improve pricing and hedging of long-term contracts. In response, a class of hybrid pricing models emerged, predominantly with applications in equity, insurance and foreign exchange markets, for instance Ballotta and Haberman (2003), Schrage and Pelsser (2004), van Haastrecht, Lord, Pelsser, and Schrage (2009), van Haastrecht and Pelsser (2011) and Grzelak, Oosterlee, and van Weeren (2012). In particular, van Haastrecht et al. (2009), van Haastrecht and Pelsser (2011) and Grzelak et al. (2012) discuss numerical solutions of models combining the stochastic volatility Schöbel and Zhu (1999) model and the stochastic interest rate Hull and White (1990) model, with full correlations between the underlying processes. van Haastrecht et al. (2009) apply the hybrid model to the valuation of insurance options with long-term equity or foreign exchange (FX) exposure. Equity and FX markets have been extensively studied with hybrid pricing

---

<sup>3</sup>The average daily open interest is calculated by summing all the open interest at the end of each trading day with maturities more than or equal to 24 months (for  $\geq 24$ ) or 36 months (for  $\geq 36$ ) over a given year and then dividing this sum by the number of trading days in that year.

models, yet there is limited literature on commodity markets. This paper aims to contribute to closing this gap.

One of the earliest commodity derivatives models is the Gibson and Schwartz (1990) model with stochastic convenience yield and constant interest rates, examining pricing performance of short-dated futures contracts. Schwartz (1997) confirms the importance of stochastic interest rates for pricing longer maturity crude oil futures contracts, yet with deterministic volatility specifications and by fitting the model to crude oil futures prices only. Hilliard and Reis (1998) propose a commodity pricing model with stochastic convenience yields, stochastic interest rates and jumps in spot prices and investigate how these features impact futures, forwards and futures options. By modelling the short-run deviations and the equilibrium level, Schwartz and Smith (2000) provide a better fit to medium-term futures prices rather than short-term and long-term futures prices. None of these models include stochastic volatility, an important feature for pricing long-dated contracts, as it has been shown in Bakshi et al. (2000) and Cortazar, Gutierrez, and Ortega (2016a). A representative, more recent literature of pricing commodity contingent claims with spot price models includes Cortazar and Schwartz (2003), Casassus and Collin-Dufresne (2005), Geman (2005), Geman and Nguyen (2005), Cortazar and Naranjo (2006) (futures prices), Dempster, Medova, and Tang (2008) (spread options) and Fusai, Marena, and Roncoroni (2008) (Asian-style options). These spot price models for commodity derivatives capture efficiently characteristics of the different commodity markets, for instance, the mean reversion of agricultural commodities and the seasonality of energy commodities. However, they have the following limitations: a) to evaluate derivative prices, one needs to model and estimate the commodity convenience yield, an unobservable variable; b) unspanned stochastic volatility cannot be easily integrated in spot price models; c) it is difficult to obtain consistency with the term structure of the forward commodity curve and d) in general, they do not provide closed-form solutions for commodity derivatives.

Futures price models do not retain these limitations. Trolle and Schwartz (2009) introduced a forward-price model for commodity derivatives and by using crude oil futures and futures options, they empirically demonstrate the existence of unspanned volatility components in the crude oil derivatives market. By fitting their model to a longer dataset of crude

oil futures and options, Chiarella, Kang, Nikitopoulos, and Tô (2013) consider a commodity pricing model within the Heath, Jarrow, and Morton (1992) framework and demonstrate that a hump-shaped crude oil futures volatility structure provides a better fit to futures and option prices and improves hedging performance. Pilz and Schlögl (2013) model commodity forward prices with stochastic interest rates driven by a multi-currency LIBOR Market Model and achieve a consistent cross-sectional calibration (i.e., single day) of the model to at-the-money market data for interest rate options, commodity options and historically estimated correlations.<sup>4</sup> Cortazar et al. (2016a) investigate the pricing performance of different models on commodity prices, namely crude oil, gold and copper. The constant volatility model fits futures prices better, but fitting to option prices improves significantly when a stochastic volatility model is considered. Chiarella, Kang, Nikitopoulos, and Tô (2016) present an alternative approach to study the return-volatility relationship in commodity futures markets and analyse this relation in the crude oil futures markets and the gold futures markets. However, most of these studies assume deterministic interest rates, thus they may not be suitable for the evaluation of long-dated contracts.

We make both theoretical and empirical contributions to the literature. On theoretical grounds, we develop a multi-dimensional stochastic volatility Heath et al. (1992) type model that accommodates stochastic interest rates. The model features a full correlation structure between the underlying variables and matches the entire initial forward commodity curve by construction. The stochastic interest rate process is modelled by a Hull and White (1990) process<sup>5</sup> and the volatility is modelled by an Ornstein-Uhlenbeck process. It is a stochastic volatility model that allows for multiple volatility factors with flexible volatility structures. Empirical evidence in the crude oil futures market demonstrates that exponential decaying or hump-shaped volatility curves are typical structures of its volatility factors, see Chiarella et al. (2013), and it features unspanned volatility, see Trolle and Schwartz (2009).

---

<sup>4</sup>Cross-sectional calibration to market data for interest rate and commodity option volatility surfaces is treated in a stochastic volatility extension of the LIBOR Market Model in Karlsson, Pilz, and Schlögl (2016).

<sup>5</sup>More specifically, it is modelled by the multidimensional Gauss/Markov generalisation of the Hull and White (1990) model. Other interest rate specifications can be considered, for instance Cox, Ingersoll, and Ross (1985), at the cost of losing the tractability of the option pricing equation, thus relying on numerical methods to solve it.

This class of models has an affine term structure representation that leads to quasi-analytical European vanilla futures option pricing equations. Furthermore, it allows us to model directly multiple futures prices simultaneously.<sup>6</sup> In contrast, spot price models require the modelling of the spot price and the interest rates, as well as making assumptions about the convenience yield to be able to specify the futures prices. The proposed approach directly models the full term structure of futures prices and provides tractable prices for futures options. Thus, our model can be estimated by fitting to both futures prices and option prices.<sup>7</sup>

From an empirical point of view, our main contributions are threefold. Firstly, we estimate the parameters of the proposed models, from historical time series of both crude oil futures prices and crude oil futures option prices (since this is the most active commodity derivatives market), by using an extended Kalman filter maximum log-likelihood methodology. Due to the large number of parameters, the estimation process is treated in three stages. In stage one, we estimate the parameters of the interest rate models by fitting the implied yields to US Treasury yield rates. In stage two, we estimate the parameters of the stochastic volatility futures price model using prices of crude oil derivatives. In stage three, we run a sensitivity analysis to determine the impact of correlations between the underlying variables on long-dated crude oil option prices, and estimate the correlation parameters. We find evidence for a negative correlation between the futures price process and the interest rate process, especially over periods of high interest rate volatility, and an insignificant impact of the correlation between stochastic volatility and interest rates on option prices.

Secondly, we evaluate the contribution of stochastic interest rate models to improving the pricing performance on long-dated crude oil derivatives, compared to models with deterministic interest rate specifications, in-sample as well as out-of-sample. To assess the impact of stochastic interest rates under different market conditions, we consider three periods; *August*

---

<sup>6</sup>Note that it is the presence of convenience yields which permits us to specify directly the prices of futures to several maturities, simultaneously on the same underlying, without introducing inconsistency to the model. We focus on modelling observable quantities and unobservables such as convenience yields are implied endogenously by the model in a consistent manner.

<sup>7</sup>One should also note that any seasonality in the spot price will be reflected in the term structure of futures prices, which the model fits by construction.

*2005 – July 2007* which was characterised by relatively volatile interest rates, *July 2007 – July 2009* which covers the Global Financial Crisis (GFC) where the markets experienced substantial turmoil, and *2011–2012* which exhibited very low interest rates and a very stable interest rate market. Numerical investigations reveal that the volatility of interest rates plays an important role in the pricing of long-dated commodity derivatives, thus the selection of periods studied is made based on different volatility levels. By comparing the pricing errors of the two models (deterministic interest rates against stochastic interest rates), we find that the stochastic interest rates counterpart typically performs better compared to the deterministic interest rates counterpart, an effect that is more pronounced as the maturity of the crude oil futures options increases. These results are more evident during periods of high volatility of interest rates.

Thirdly, we investigate the dimensionality of the model required to provide better pricing performance on long-dated crude oil derivatives contracts. We find that three-dimensional models provide better fit to futures prices of all maturities compared to two-dimensional models. However, for long-dated crude oil option prices, three-dimensional models matter only when the markets are volatile, see also Schwartz and Smith (2000) and Cortazar et al. (2016a) for similar conclusions.

All in all, this leads to a quite high-dimensional model with a correspondingly large number of parameters. This is a consequence of the objective to investigate the additional contribution of including stochastic interest rates beyond stochastic volatility in a model of commodity derivatives — given the literature, it is accepted that the latter is important, and therefore must be included in an analysis investigating the contribution of the former. In particular in light of this high dimensionality, we base our conclusions on out-of-sample performance, rather than in sample where overfitting is a possibility.

The paper is structured as follows. Section 2 presents a term structure model for pricing commodity derivatives incorporating stochastic volatility and stochastic interest rates. Section 3 describes the crude oil derivative data used in the empirical analysis and provides the details of the estimation methodology. Section 4 presents the estimation results and discusses the empirical findings of pricing long-dated crude oil derivatives. Section 5 concludes.

## 2. Commodity derivatives models and stochastic interest rates

We develop a futures price model for commodity derivatives incorporating stochastic volatility and stochastic interest rates. The model can be easily adjusted to allow for deterministic interest rate specifications. Using these models, we assess the contribution of stochastic interest rates beyond stochastic volatility when pricing long-dated commodity derivatives. We first present the stochastic volatility–stochastic interest rate model and then we specify a deterministic interest rate model.

### 2.1. Model setup

We consider a filtered probability space  $(\Omega, \mathcal{F}_T, \mathbb{F}, \mathbb{P}), T \in [0, \infty)$  satisfying the usual conditions<sup>8</sup>. Here  $\Omega$  is the state-space,  $\mathbb{F} = \{\mathcal{F}_t\}_{t \in [0, T]}$  is a set of  $\sigma$ -algebras representing measurable events and  $\mathbb{P}$  is the historical (real-world) probability measure. We denote by  $\mathbb{Q}$  the risk-neutral probability measure equivalent to  $\mathbb{P}$ , under which all asset prices discounted by the continuously compounded savings account are martingales.<sup>9</sup> We introduce  $\sigma = \{\sigma(t); t \in [0, T]\}$ , an  $n$ -dimensional stochastic volatility process modelling the uncertainty in the commodity market, where we further assume that the components of the volatility process  $\sigma(t) = \{\sigma_i(t), \dots, \sigma_n(t)\}$  follow the dynamics<sup>10</sup>

$$d\sigma_i(t) = \kappa_i(\bar{\sigma}_i - \sigma_i(t))dt + \gamma_i dW_i^\sigma(t), \quad (1)$$

where  $W^\sigma(t) = \{W_1^\sigma(t), \dots, W_n^\sigma(t)\}$  is an  $n$ -dimensional vector of independent Wiener processes under  $\mathbb{Q}$  and for all  $t \in [0, T]$ ,  $\kappa_1, \dots, \kappa_n$ ,  $\bar{\sigma}_1, \dots, \bar{\sigma}_n$ , and  $\gamma_1, \dots, \gamma_n$  are constants.

We let  $F(t, T, \sigma(t))$  be the futures price of the commodity at time  $t \geq 0$ , for delivery at time  $T \in [t, \infty)$ , highlighting the dependence of  $F$  on the current state of the stochas-

---

<sup>8</sup>The usual conditions satisfied by a filtered complete probability space are: (a)  $\mathcal{F}_0$  contains all the  $\mathbb{P}$ -null sets of  $\mathcal{F}$  and (b) the filtration is right continuous

<sup>9</sup>Under the assumptions on the specific stochastic dynamics which follow, the existence of  $\mathbb{Q}$  is equivalent to the absence of arbitrage by results standard in the literature.

<sup>10</sup>Note that there is no mathematical issue with the fact that this implies that  $\sigma_i(t)$  may become negative — this simply changes the sign of the impact of a change in  $W_i^x(t)$  on futures prices. This observation is not new: A number of papers in the literature use an Ornstein/Uhlenbeck process such as this to model stochastic volatility, including Stein and Stein (1991), Schöbel and Zhu (1999), Lipton and Sepp (2008) and van Haastrecht et al. (2009).



tic volatility process  $\sigma(t) \in \mathfrak{R}^n$ . Under the assumed dynamics (3) below, paths of futures prices are almost surely continuous, thus the spot price at time  $T$  of the underlying commodity  $S(T)$  can be obtained by taking the limit of the futures price as  $t \rightarrow T$ , i.e.  $S(T) = \lim_{t \rightarrow T} F(t, T, \sigma(t))$ ,  $t \in [0, T]$ , where this limit exists almost surely. We denote  $r = \{r(t); t \in [0, T]\}$  the stochastic process of the instantaneous short-rate of interest. By using no-arbitrage arguments, Cox, Ingersoll, and Ross (1981) demonstrate that the futures price process  $F(t, T, \sigma(t))$  is a martingale under  $\mathbb{Q}$ , i.e.

$$F(t, T, \sigma(t)) = \mathbb{E}^{\mathbb{Q}}[S(T, \sigma(T)) | \mathcal{F}_t].$$

Under the general Heath et al. (1992) framework, the term structure of the futures prices is infinite-dimensional, representing futures prices across a continuum of maturities. At this level of generality, the stochastic dynamics are Markovian only in an infinite dimensional state-space. We instead proceed along the lines of Chiarella and Kwon (2003), Björk, Landén, and Svensson (2004) and Björk, Blix, and Landén (2006), restricting ourselves to a system that can admit finite dimensional realisations, when the functional form of the futures price volatilities  $\sigma_i^F(t, T, \sigma(t))$  is specified as follows:

$$\sigma_i^F(t, T, \sigma(t)) = (\xi_{0i} + \xi_i(T - t))e^{-\eta_i(T-t)}\sigma_i(t) \quad (2)$$

with  $\xi_{0i}$ ,  $\xi_i$ , and  $\eta_i \in \mathfrak{R}$  for all  $i \in \{1, \dots, n\}$ . Under  $\mathbb{Q}$ , the commodity futures price processes are then given by the driftless stochastic differential equations of the form:

$$\frac{dF(t, T, \sigma(t))}{F(t, T, \sigma(t))} = \sum_{i=1}^n \sigma_i^F(t, T, \sigma(t)) dW_i^x(t), \quad (3)$$

i.e.  $\sigma_i^F(t, T, \sigma(t))$  are the  $\mathbb{F}$ -adapted futures price volatility processes for all  $T > t$  and  $W^x(t) = \{W_1^x(t), \dots, W_n^x(t)\}$  is an  $n$ -dimensional vector of independent Wiener processes under  $\mathbb{Q}$  (in our empirical analysis, we will consider the case  $n = 2$  and  $n = 3$ ).

These volatility specifications are important for two reasons: Firstly, they allow for a variety of volatility structures, such as exponentially decaying and hump-shaped, which are typical volatility structures in the commodity market, see Trolle and Schwartz (2009) and

Chiarella et al. (2013). Secondly, they lead to tractable solutions for futures prices as the system now admits finite-dimensional realisations.

The instantaneous short rate of interest  $r(t)$  is modelled as<sup>11</sup>

$$r(t) = \bar{r}(t) + \sum_{j=1}^N r_j(t), \quad (4)$$

$$dr_j(t) = -\lambda_j(t)r_j(t)dt + \theta_j dW_j^r(t).$$

where for all  $t \in [0, T]$ ,  $\theta_1, \dots, \theta_N$  are constants, and  $\{\lambda_i(t)\}_{i=1, \dots, N}$  and  $\bar{r}(t)$  are deterministic functions of time  $t$ .<sup>12</sup> In our empirical analysis we will consider  $N = 1, 2$  and  $3$ .

We also make the following assumptions on the correlation structure of the associated

---

<sup>11</sup>As written here, this is a slightly generalised, multifactor version of the Hull and White (1990) model. If the  $\lambda_j : \mathfrak{R}_+ \rightarrow \mathfrak{R}$  are locally bounded deterministic functions, the stochastic differential equations for each of the factors  $r_j$  each have a solution and the distribution of the factors is Gaussian (see e.g. Section 12.3 of Musiela and Rutkowski (1997)). In the empirical analysis below, the  $\lambda_j$  are taken to be constant.

<sup>12</sup>Under the measure  $\mathbb{Q}$ , for a zero coupon bond  $B(0, T)$ , i.e. observed at the initial time 0 and maturing at a time  $T$ , it must hold that

$$\begin{aligned} B(0, T) &= \mathbb{E}^{\mathbb{Q}} \left[ \exp \left\{ - \int_0^T r(t) dt \right\} \right] \\ &= \exp \left\{ - \int_0^T \bar{r}(t) dt \right\} \mathbb{E}^{\mathbb{Q}} \left[ \exp \left\{ - \int_0^T \sum_{j=1}^N r_j(t) dt \right\} \right] \end{aligned}$$

In practical applications where this is required, such as the relative pricing of fixed income derivatives in a model calibrated to cross-sectional (i.e., at a single point in time) data, the function  $\bar{r}(t)$  is typically defined implicitly to ensure that the model matches the initial term structure of interest at time zero exactly, i.e.

$$\bar{r}(t) = - \frac{\partial}{\partial T} \ln \frac{B(0, T)}{\mathbb{E}^{\mathbb{Q}} \left[ \exp \left\{ - \int_0^T \sum_{j=1}^N r_j(t) dt \right\} \right]}$$

where it is assumed that the market-observed zero coupon bond prices are interpolated in a differentiable manner. In time-series estimation applications such as ours below, it is neither required nor desirable to let  $\bar{r}$  depend on  $t$ , so we take it to be constant.

Wiener processes:

$$\begin{aligned}
dW_i^x(t)dW_j^\sigma(t) &= \begin{cases} \rho_i^{x\sigma} dt, & \text{if } i = j, \\ 0, & \text{otherwise} \end{cases} \\
dW_i^x(t)dW_{\hat{j}}^r(t) &= \begin{cases} \rho_i^{xr} dt, & \text{if } \hat{j} = 1, \\ 0, & \text{otherwise} \end{cases} \\
dW_i^\sigma(t)dW_{\hat{j}}^r(t) &= \begin{cases} \rho_i^{r\sigma} dt, & \text{if } \hat{j} = 1, \\ 0, & \text{otherwise} \end{cases}
\end{aligned} \tag{5}$$

for  $i \in \{1, \dots, n\}$ ,  $j \in \{1, \dots, n\}$  and  $\hat{j} \in \{1, \dots, N\}$ . The above-mentioned specifications result in unspanned stochastic volatility in the model, as in Trolle and Schwartz (2009) and Chiarella et al. (2013). More specifically, when the Wiener processes  $W_i^x(t)$  and  $W_i^\sigma(t)$  are correlated, futures contracts can be used to partially hedge the volatility risk of the derivatives, while when the Wiener processes  $W_i^x(t)$  and  $W_i^\sigma(t)$  are uncorrelated, the volatility risk of the derivatives is unhedgeable by futures contracts. Note that for modelling convenience, we assume that only the first Wiener process of the interest rate process  $W_1^r(t)$  can be correlated with the futures price process and the futures volatility process. Under these assumptions we have:

**Proposition 1.** *The volatility specifications of (2) allow the instantaneous futures prices  $F(t, T, \sigma(t))$  at time  $t$  with maturity  $T$ , satisfying the dynamics (3), to be expressed in terms of  $6n$  state variables, namely  $x_i(t), y_i(t), z_i(t), \phi_i(t), \psi_i(t)$  and  $\sigma_i(t)$ :*

$$\begin{aligned}
\log F(t, T, \sigma(t)) &= \log F(0, T, \sigma(0)) \\
&\quad - \frac{1}{2} \sum_{i=1}^n \left( \gamma_{1i}(T-t)x_i(t) + \gamma_{2i}(T-t)y_i(t) + \gamma_{3i}(T-t)z_i(t) \right) \\
&\quad + \sum_{i=1}^n \left( \beta_{1i}(T-t)\phi_i(t) + \beta_{2i}(T-t)\psi_i(t) \right), \tag{6}
\end{aligned}$$

where for  $i = 1, \dots, n$  the deterministic functions are defined as:

$$\begin{aligned}
\beta_{1i}(T-t) &= \varpi_i(T-t) = (\xi_{i0} + \xi_i(T-t))e^{-\eta_i(T-t)}, \\
\beta_{2i}(T-t) &= \xi_i e^{-\eta_i(T-t)}, \\
\gamma_{1i}(T-t) &= \beta_{1i}(T-t)^2, \\
\gamma_{2i}(T-t) &= 2\beta_{1i}(T-t)\beta_{2i}(T-t), \\
\gamma_{3i}(T-t) &= \beta_{2i}(T-t)^2,
\end{aligned}$$

and the  $5n$  state variables  $x_i, y_i, z_i, \phi_i, \psi_i$  satisfy the following dynamics:

$$\begin{aligned}
dx_i(t) &= (-2\eta_i x_i(t) + \sigma_i^2(t)) dt, \\
dy_i(t) &= (-2\eta_i y_i(t) + x_i(t)) dt, \\
dz_i(t) &= (-2\eta_i z_i(t) + 2y_i(t)) dt, \\
d\phi_i(t) &= -\eta_i \phi_i(t) dt + \sigma_i(t) dW_i^x(t), \\
d\psi_i(t) &= (-\eta_i \psi_i(t) + \phi_i(t)) dt,
\end{aligned} \tag{7}$$

subject to the initial condition  $x_i(0) = y_i(0) = z_i(0) = \phi_i(0) = \psi_i(0) = 0$ . The system also includes the  $n$  stochastic volatility processes  $\sigma_i(t)$ ,  $i \in \{1, \dots, n\}$ , see equations (1), giving a total of  $6n$  state variables.

*Proof.* Along the lines of Chiarella et al. (2013). □

The stochastic interest rate process  $r(t)$  does not affect the futures price process per se, but it matters when we consider pricing options on futures contracts.

## 2.2. Affine class transformation

The initially infinite-dimensional Markovian model is now reduced to a model with  $6n + N$  state variables. The additional  $N$  state variables are from the stochastic interest rate process specified by the equation (4). Note that this system is not affine. When the volatility specifications (2) are applied to the dynamics (6), the  $\sigma_i^2(t)$  term is not an affine transformation of  $\sigma_i(t)$ , which would be required in order for the system to admit a closed-form characteristic function of  $X(t, T) = \log F(t, T, \sigma(t))$ , see Duffie, Pan, and Singleton (2000).

By introducing a latent stochastic variable  $\nu_i(t) \triangleq \sigma_i^2(t)$  with  $\nu_t = \{\nu_1(t), \dots, \nu_n(t)\}$ , this system can be transformed into the following affine system:

$$dX(t, T) = -\frac{1}{2} \sum_{i=1}^n \beta_{1i}^2(T-t) \nu_i(t) dt + \sum_{i=1}^n \beta_{1i}(T-t) \sqrt{\nu_i(t)} dW_i^x(t),$$

where, for  $i = 1, 2, \dots, n$ ,

$$\begin{aligned} d\sigma_i(t) &= \kappa_i(\bar{\sigma}_i - \sigma_i(t))dt + \gamma_i dW_i^\sigma(t), \\ d\nu_i(t) &= 2\kappa_i(\bar{\sigma}_i \sigma_i(t) + \frac{\gamma_i^2}{2\kappa_i} - \nu_i(t))dt + 2\gamma_i \sqrt{\nu_i(t)} dW_i^\sigma(t), \end{aligned}$$

and the instantaneous short-rate process  $r(t)$  follows the dynamics (4), with the correlation structure being the same as in (5). For  $t \leq T_o \leq T$  and  $v \in \mathbb{C}$ , the  $r_1(t)$ -discounted characteristic functions of the logarithm of the futures prices  $\phi(t) \triangleq \phi(t, X(t, T), r_1(t), \nu(t), \sigma(t); v, T_o, T)$ :

$$\begin{aligned} \phi(t; v, T_o, T) &\triangleq \mathbb{E}_t^\mathbb{Q} \left[ e^{-\int_t^{T_o} r_1(u) du} \exp \{v \log F(T_o, T, \sigma(T_o))\} \right] \\ &= \mathbb{E}_t^\mathbb{Q} \left[ e^{-\int_t^{T_o} r_1(u) du} \exp \{v X(T_o, T)\} \right] \end{aligned} \quad (8)$$

can be expressed as:

$$\begin{aligned} \phi(t; v, T_o, T) &= \exp \{A(t; v, T_o) + B(t; v, T_o)X(t, T) + C(t; v, T_o)r_1(t) \\ &\quad + \sum_{i=1}^n D_i(t; v, T_o)\nu_i(t) + \sum_{i=1}^n E_i(t; v, T_o)\sigma_i(t)\}. \end{aligned} \quad (9)$$

**Lemma 1.** *The functions  $A(t; v, T_o), B(t; v, T_o), C(t; v, T_o), D_i(t; v, T_o)$  and  $E_i(t; v, T_o)$  in*

equation (9) satisfy the following complex-valued Ricatti ordinary differential equations:

$$\begin{aligned}
\frac{\partial B}{\partial t} &= 0, \\
\frac{\partial C}{\partial t} &= \lambda_1 C + 1, \\
\frac{\partial D_i}{\partial t} &= -\frac{1}{2}\beta_{1i}^2(T-t)(B-1)B - 2(\rho_i^{x\sigma}\beta_{1i}(T-t)\gamma_i B - \kappa_i)D_i - 2\gamma_i^2 D_i^2, \\
\frac{\partial E_i}{\partial t} &= -2\bar{\sigma}_i\kappa_i D_i - \rho_i^{xr}\theta_1\beta_{1i}(T-t)BC - 2\rho_i^{r\sigma}\theta_1\gamma_i CD \\
&\quad - (2\gamma_i^2 D_i - \kappa_i + \rho_i^{x\sigma}\beta_{1i}(T-t)\gamma_i B)E_i, \\
\frac{\partial A}{\partial t} &= -\frac{1}{2}\theta_1^2 C^2 - \sum_{i=1}^n \gamma_i^2 D_i - \sum_{i=1}^n (\kappa_i\bar{\sigma}_i + \frac{1}{2}\gamma_i^2 E_i + \rho_i^{r\sigma}\theta_1\gamma_i C)E_i,
\end{aligned} \tag{10}$$

where  $i \in \{1, \dots, n\}$ , subject to the terminal condition  $\phi(T_o) = e^{vX(T_o, T)}$ .

*Proof.* Along the lines of Duffie et al. (2000). □

In the next section, we present the quasi-analytical pricing formulae for European vanilla options on futures resulting from this affine transformation.

### 2.3. Pricing of European Option on Futures

We denote by  $\text{Call}(t, F(t, T, \sigma(t)); T_o)$  and  $\text{Put}(t, F(t, T, \sigma(t)); T_o)$  the price of the European style call and put option, respectively, with maturity  $T_o$  and strike  $K$  on the futures price  $F(t, T, \sigma(t))$  maturing at time  $T$ . The price of a call option can be expressed as the discounted expected payoff under the risk-neutral measure:

$$\text{Call}(t, F(t, T, \sigma(t)); T_o) = \mathbb{E}_t^{\mathbb{Q}} \left( e^{-\int_t^{T_o} r(s) ds} (e^{X(T_o, T)} - K)^+ \right). \tag{11}$$

By using the Fourier inversion technique, Duffie et al. (2000) provide a semi-analytical formula for the price of European-style vanilla options in the class of affine term structure models. With a slight modification of the pricing equation in Duffie et al. (2000), equation

(11) can be expressed as:

$$\begin{aligned} \text{Call}(t, F(t, T, \sigma(t)); T_o) &= e^{-\int_t^{T_o} \bar{r}(s) ds} \prod_{i=2}^N \mathbb{E}_t^{\mathbb{Q}} [e^{-\int_t^{T_o} r_i(s) ds}] \times \\ & [G_{1,-1}(-\log K) - KG_{0,-1}(-\log K)], \end{aligned} \quad (12)$$

where

$$G_{a,b}(y) = \frac{\phi(t; a, T_o, T)}{2} - \frac{1}{\pi} \int_0^\infty \frac{\text{Im}[\phi(t; a + \mathbf{i}bu, T_o, T)e^{-\mathbf{i}uy}]}{u} du. \quad (13)$$

Note that  $\mathbf{i}^2 = -1$  and  $\text{Im}(x + \mathbf{i}y) = y$ . Note that the product starts at  $i = 2$  because the equation (12) is specific to the correlation structure (5) that allows only the first factor of the interest rate process to be correlated with the futures price process. For European put options, the discounted expected payoff is:

$$\begin{aligned} \text{Put}(t, F(t, T, \sigma(t)); T_o) &= \mathbb{E}_t^{\mathbb{Q}} \left( e^{-\int_t^{T_o} r(s) ds} (K - e^{X(T_o, T)})^+ \right) \\ &= e^{-\int_t^{T_o} \bar{r}(s) ds} \prod_{i=2}^N \mathbb{E}_t^{\mathbb{Q}} [e^{-\int_t^{T_o} r_i(s) ds}] \times \\ & [KG_{0,1}(\log K) - G_{1,1}(\log K)]. \end{aligned} \quad (14)$$

Quasi-analytical option pricing formulae greatly facilitate model estimation and calibration, which we exploit in the next section where the model is estimated using historical data from the crude oil derivatives market. Since the futures prices observed from the market are under the historical measure, for estimation applications, we need to account for the market price of risk and the market price of volatility risk, namely  $\varpi_i$  and  $\varpi_i^\sigma$ , respectively. Introducing  $\mathbb{W}_i(t)$  and  $\mathbb{W}_i^\sigma(t)$ , standard Wiener processes under the physical measure  $\mathbb{P}$ , we have

$$\begin{aligned} d\mathbb{W}_i(t) &= dW_i^x(t) - \varpi_i \sqrt{\nu_i(t)} dt, \\ d\mathbb{W}_i^\sigma(t) &= dW_i^\sigma(t) - \varpi_i^\sigma \sqrt{\nu_i(t)} dt. \end{aligned}$$

For comparison purposes, we also consider a model with deterministic interest rates, which is described next.

#### 2.4. The deterministic interest rate model

The dynamics of the futures price process and the stochastic volatility process remain the same as specified in equations (6) and (1), but the stochastic interest rate process is replaced by a deterministic discount function. This discount function  $P_{NS}(t, T)$  is obtained by using the Nelson and Siegel (1987) functional form of the time- $t$  instantaneous forward rate as follows:

$$f_{NS}(t, T) = b_0 + b_1 e^{-a(T-t)} + b_2 a(T-t) e^{-a(T-t)}$$

where  $a, b_0, b_1$  and  $b_2$  are constant parameters to be determined. The yield to maturity is then expressed as:

$$y_{NS}(t, T) = \frac{\int_t^T f_{NS}(t, u) du}{T-t} = b_0 + \frac{(b_1 + b_2)(1 - e^{-a(T-t)})}{a(T-t)} - b_2 e^{-a(T-t)}. \quad (15)$$

The discount function for the option with maturity  $T$  is simply  $P_{NS}(t, T) = e^{-y_{NS}(t, T)(T-t)}$ .

We next conduct an empirical study using crude oil derivatives. The aim of the study is to gauge the impact of stochastic interest rates beyond stochastic volatility on pricing long-dated crude oil derivatives.

### 3. Description of data and estimation methodology

Using the models developed in Section 2, we conduct a sensitivity analysis of the impact the different interest rate parameters have on commodity derivative prices. Since the objective is to explore parameter sensitivities to suggest fruitful directions of inquiry in the empirical analysis which follows in Section 4, the base parameters in this exploration are chosen in a deliberately stylised fashion, so as to better illustrate the relevant effects (leaving the analysis on empirically realistic parameter constellations to Section 4).<sup>13</sup> The analysis reveals that the correlation between the stochastic interest rates and the stochastic futures price process has noticeable impact on prices of long-dated options, but remains negligible for short-dated options, see Figure 2. Furthermore, Figure 3 reveals that as the volatility

---

<sup>13</sup>For this sensitivity analysis, we consider a one-dimensional version of the model for pricing ATM options with the following parameter values;  $\bar{r} = 0.05$ ,  $\lambda = 4$ ,  $\theta = 0.05$ ,  $\kappa = 4$ ,  $\bar{\sigma} = 0.5$ ,  $\gamma = 0.6$ ,  $K = 100$ .



of interest rates increases, the value of the option increases with the impact being more pronounced for longer-maturity options, while the long-term level of the interest rates does not impact option prices in a similar manner.

Accordingly, in order to assess the contribution of stochastic interest rates on long-dated crude oil derivative prices, we select three two-year periods to estimate the model parameters, each period with marked differences in terms of interest rate levels and, most importantly, volatilities. After investigations, based on descriptive statistics (see Table 1) and properties of the interest rate process and futures price process (see Section 4.1), we have identified three two-year periods that capture three different environments of the interest rates markets. The first period is from 1<sup>st</sup> August 2005 to 31<sup>st</sup> July 2007 and it represents a period of relatively high levels of interest rates (over 4.6%, see Table 1) with average levels of interest rate volatility. The second period is from 1<sup>st</sup> July 2007 to 30<sup>th</sup> June 2009. This period spreads over the GFC, where the level of interest rates has almost halved in the course of enormously volatile market conditions. This period exhibits the highest volatility of interest rates over the sample. The third period is from 1<sup>st</sup> January 2011 to 31<sup>st</sup> December 2012, a period of extremely low interest rates featuring also low volatilities (interest rates were below 0.5% for maturities under three years, see Table 1). This period is one of the lowest interest rate volatility periods in the sample.<sup>14</sup> These three periods can also be considered as pre-GFC, during GFC and post-GFC sample periods.

### *3.1. Data description*

#### *3.1.1. Interest rate data*

We use the US Treasury yield rates<sup>15</sup> as the proxy to estimate the parameters in the interest rate process in our model as well as to convert the prices of American options to European options, as required when we consider the crude oil options. The reason for this choice over other rates such as LIBOR is that the options in the dataset are exchange-traded options, hence there is no counterparty credit risk involved. There are over ten different

---

<sup>14</sup>From 2013 onwards, the behaviour of interest rates in terms of level and volatility falls between the extremes covered by the three periods studied here (see Figure 4), and does not add any new quality relative to the results reported in our empirical analysis (and is therefore omitted).

<sup>15</sup>Data were obtained from [www.treasury.gov](http://www.treasury.gov).

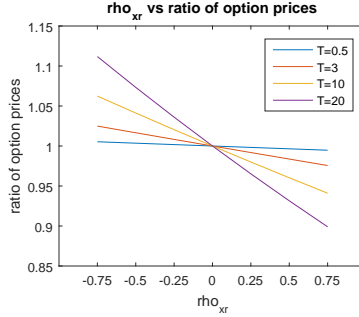


Figure 2: **Impact of the correlation between futures prices and interest rates on call option prices.** The plot shows the ratio of option prices (between prices from a model with full correlation structure and prices from a restricted model with  $\rho^{r\sigma} = 0$ ) for a range of correlation coefficient parameter values and for four different maturities; 0.5, 3, 10 and 20 years.

maturities in the dataset and we choose only four sets of yield rates (the 1-year, 2-year, 3-year and 5-year yields) that best match the maturities of the options used in the analysis. Figure 4 displays the evolution of the US Treasury yields from 2005 to 2016, which includes the sample periods used in our empirical study.

Summary statistics of the yields over the three sample periods considered in our analysis are presented in Table 1. It is evident that the interest rate markets behaved very differently over these three periods. The mean levels of interest rates have reduced from around 4.5% in 2005 to around 2.5% in 2008 and down to 0.5% in 2011. Yet, the market experienced high volatility over the GFC, with lower volatility levels in the pre-GFC period and extremely low volatility in the post-GFC period. Note that the standard deviation of the 5-year bond yields estimated in the period between January 2011 and December 2012 (0.539%) is almost double the standard deviation (0.288%) of the corresponding yield in the period August 2005 to July 2007. The reason for this seemingly higher volatility is that the 5-year yield at the beginning of 2011 is around 2% and it increases to 2.4% in early February 2011 and then it quickly plummets to less than 1% in the beginning of 2011. To better quantify the volatility of the interest rates during those three periods, we also present the linearly, as well as the nonlinearly detrended standard deviation of the interest rates. The linearity and nonlinearity are removed by subtracting a least-squares polynomial fit of degree 1 and degree 2 respectively. We observe that the nonlinearly detrended standard deviation of the 5-year yield is lower during the post-GFC period.

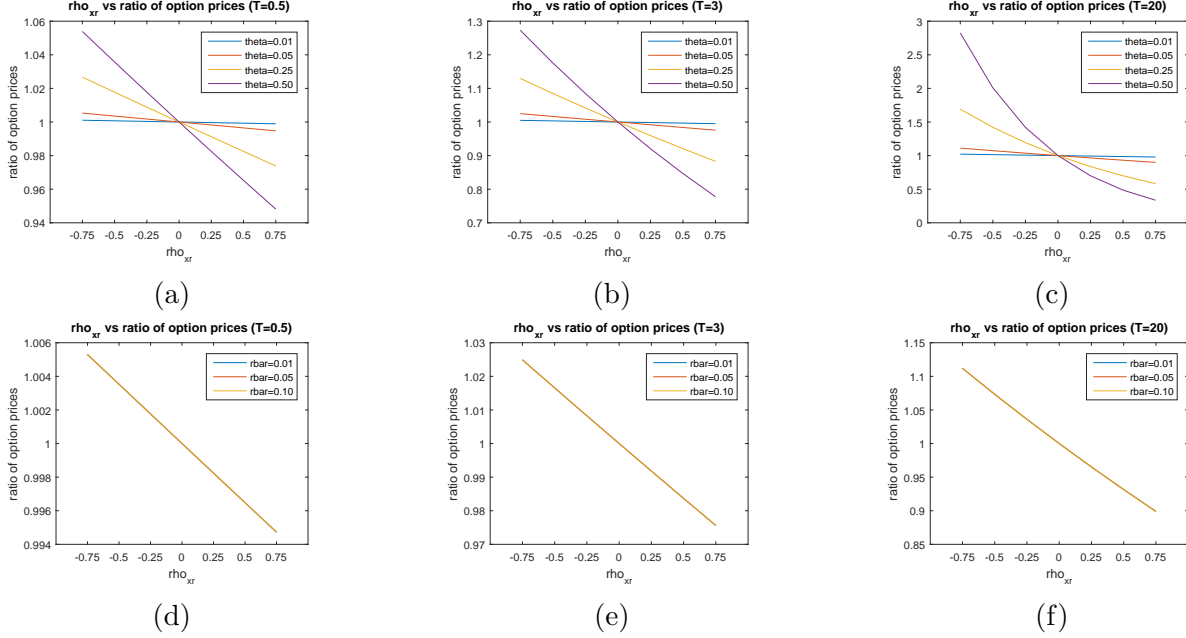


Figure 3: **Impact of the volatility and long term levels of interest rates on call option prices.** The plot displays ratios of option prices (between prices from a model with full correlation structure and prices from a restricted model with  $\rho^{xr} = 0$ ) for varying levels of volatility of interest rates  $\theta = 0.01, 0.05, 0.25, 0.50$  (top three graphs) and ratios of option prices for varying long term levels of interest rates  $\{\bar{r}, r_0\} = \{0.01, 0.01\}, \{0.05, 0.05\}, \{0.10, 0.10\}$  (bottom three graphs) and for three different maturities  $T = 0.5, T = 3$  and  $T = 20$ .

### 3.1.2. Crude oil derivative data

We use Light Sweet Crude Oil (WTI) futures and options traded on the NYMEX<sup>16</sup>, which is one of the richest datasets available on commodity derivatives. The average daily number of actively traded futures contracts has increased from 34 in 2005 to 59 in 2012. The futures dataset has 145,805 lines of data and the options dataset has close to 5 million lines of data. Due to the enormous amount of data, for estimation purposes we make a selection of contracts based on their liquidity and we use the open interest of the futures as the proxy of its liquidity.<sup>17</sup> Figure 5 shows the open interest of the futures contracts by time-to-maturity for the first nine months and then the open interest of the futures contracts with maturities of more than one year by calendar month. It is clear that liquidity is mainly

<sup>16</sup>The database has been provided by CME.

<sup>17</sup>Liquidity is generally very low for crude oil futures and options contracts with less than 14 days to expiration. The liquidity of contracts with more than 14 days to expiration increases significantly. Furthermore, prices for contracts with very short maturities may be erratic. Thus in our analysis, we use only contracts with more than 14 days to expiration.

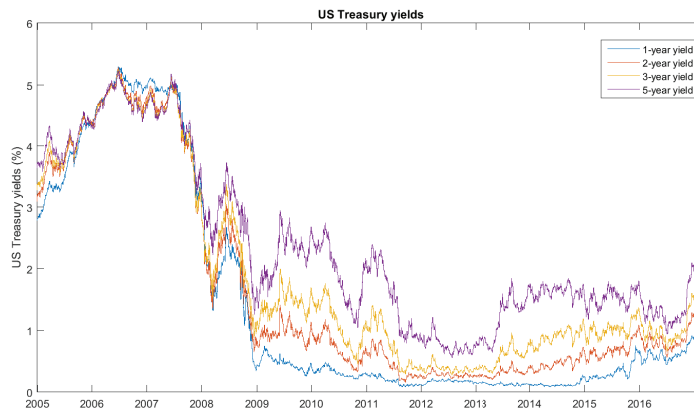


Figure 4: **US Treasury yields.** The figure displays the US Treasury yields with 1-year, 2-year, 3-year and 5-year maturity from January 2005 to December 2016.

concentrated on short-maturity contracts and on the December contracts. Thus we select the first seven monthly contracts, then the next three contracts with maturity either in March, June, September or December and then all contracts with maturity in December. We also filter out abnormalities such as futures prices of zero and zero open interest. Thus on a daily basis, we use around 10 – 15 futures contracts for the period 2005–2007, 2007–2009, and 15 – 17 futures contracts for the period 2011–2012, extending our dataset of futures maturities to eight years. For the crude oil futures option dataset, we select the options with the underlying futures contracts being the ones used in the crude oil futures dataset. That is, we take options on the first seven monthly futures contracts and the next three contracts with maturity in either March, June, September, and all December contracts with maturities of up to 5 years. For each option maturity, we consider six moneyness intervals, 0.86–0.905, 0.905–0.955, 0.955–1.005, 1.005–1.055, 1.055–1.105 and 1.105–1.15. We define moneyness as the option strike divided by the price of the underlying futures contract. In each moneyness interval, we choose either call or put options, in order to use only the out-of-the-money (OTM) and at-the-money (ATM) options that are closest to the interval mean. In order to reduce the computational overhead, we select the OTM and ATM options, as they are generally more liquid. Besides that, the OTM options have lower early exercise approximation errors. On a daily basis we use around 50 – 77 options for the period 2005 – 2007, around 60 – 80 options for the period 2007 – 2009 and 77 – 95 options for the

Period 1: August 2005 - July 2007				
Maturity in years	1	2	3	5
Mean	4.773%	4.682%	4.648%	4.640%
Standard Deviation	0.389%	0.309%	0.297%	0.288%
Detrended S.D.	0.262%	0.250%	0.251%	0.247%
Nonlinear Detrended S.D.	0.115%	0.177%	0.198%	0.210%
Excess Kurtosis	0.411	0.268	0.012	-0.252
Skewness	-1.222	-0.795	-0.557	-0.334
Period 2: July 2007 - June 2009				
Maturity in years	1	2	3	5
Mean	2.075%	2.234%	2.462%	2.944%
Standard Deviation	1.381%	1.194%	1.074%	0.913%
Detrended S.D.	0.504%	0.522%	0.541%	0.499%
Nonlinear Detrended S.D.	0.422%	0.453%	0.472%	0.452%
Excess Kurtosis	-0.774	-0.696	-0.624	-0.604
Skewness	0.556	0.585	0.599	0.303
Period 3: January 2011 - December 2012				
Maturity in years	1	2	3	5
Mean	0.178%	0.363%	0.565%	1.140%
Standard Deviation	0.052%	0.173%	0.310%	0.539%
Detrended S.D.	0.049%	0.122%	0.194%	0.264%
Nonlinear Detrended S.D.	0.036%	0.086%	0.131%	0.172%
Excess Kurtosis	-0.026	0.783	0.343	-0.520
Skewness	0.450	1.415	1.324	0.991

Table 1: **Descriptive statistics of US Treasury yields.** The table displays the descriptive statistics for the 1-, 2-, 3- and 5-year US Treasury yields over three periods; August 2005 – July 2007, July 2007–June 2009 and January 2011 – December 2012.

period 2011 – 2012.

These are American-style option prices, thus for the conversion to European-style option prices, we use the Barone-Adesi and Whaley (1987) approximation method to back out the corresponding implied volatilities, see Trolle and Schwartz (2009) for similar treatment. The Barone-Adesi and Whaley (1987) approximation assumes a constant risk-free interest rate, thus it can perform well for pricing short-dated contracts but it can potentially bring some bias when pricing long-dated contracts. To reduce this bias, we compute this “constant” risk-free interest rate parameter from bond prices that are inferred from market information. More specifically, by using the Kalman filter and maximum likelihood, we estimate our multi-factor stochastic interest rate model by fitting to US Treasury Bills data. Then by

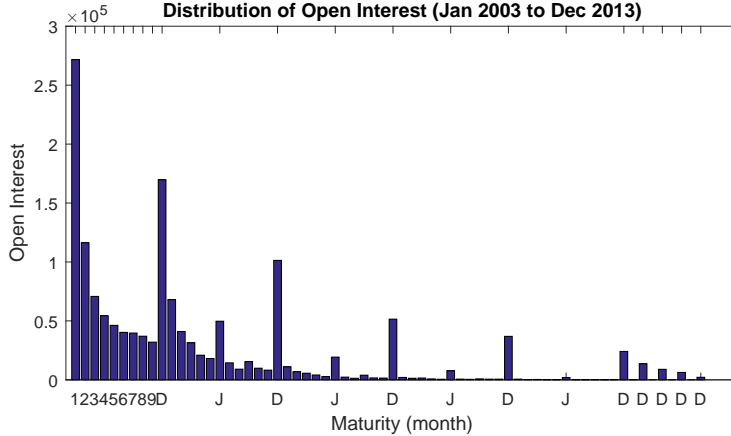


Figure 5: **Liquidity of crude oil futures contracts by calendar month.** The plot shows the liquidity of the first nine months from the trading day as well as December and June contracts in the following years. It shows that futures contracts with next-month delivery date are the most liquid and liquidity gradually decreases over the following months. Futures contracts with maturities in December are very liquid even after a few years and June contracts are moderately liquid. Contracts with maturities more than two years are very illiquid in other months.

using the calibrated model, we calculate zero-coupon bonds, which are used to compute the “constant” risk-free interest rate.<sup>18</sup>

### 3.2. Estimation method

Several methodologies have been proposed in the literature to estimate the parameters of stochastic models, such as efficient method of moments, see Gallant, Hsieh, and Tauchen (1997), maximum likelihood estimation, see Chen and Scott (1993), and the Kalman filter method. Duffee and Stanton (2012b) perform an extensive analysis and comparison of these methods and conclude that the Kalman filter is the best method among these three. In this paper, we adopt the Kalman filter methodology to estimate the parameters of our model. For the purpose of parameter estimation, we let  $\bar{r}(t)$ ,  $\lambda_i(t)$  and  $\theta_i(t)$  be constants for all  $i$ , i.e.  $\bar{r}(t) = \bar{r}$ ,  $\lambda_i(t) = \lambda_i$  and  $\theta_i(t) = \theta_i, \forall i$ . With these time-varying functions taken as constants

<sup>18</sup>Despite this treatment, the approximation can potentially lead to mis-pricing and bias. This may well reduce the improvement stochastic interest rate models bring over the deterministic counterparts for long-dated contracts in our analysis, but does not detract from our results: Even using this approximation, we find that stochastic interest rate models price more accurately long-dated contracts over periods of high interest rate volatility. Further investigations on the bias resulting from applying the Barone-Adesi and Whaley (1987) approximation to long-dated contracts, and additional improvements resulting from mitigating this bias, are left for further research.

the option pricing formula (12) is reduced to:

$$\begin{aligned} \text{Call}(t, F(t, T, \sigma(t)); T_o) &= e^{-\bar{r}(T_o-t)} \prod_{i=2}^N \exp(-A_i(T_o-t)r_i(t) + D_i(T_o-t)) \times \\ & [G_{1,-1}(-\log K) - KG_{0,-1}(-\log K)], \end{aligned} \quad (16)$$

where

$$\begin{aligned} A_i(\tau) &= \frac{1 - e^{-\lambda_i \tau}}{\lambda_i}, \\ D_i(\tau) &= \left( -\frac{\theta_i^2}{2\lambda_i^2} \right) (A_i(\tau) - \tau) - \frac{\theta_i^2 A_i(\tau)^2}{4\lambda_i}. \end{aligned}$$

There are  $34 + 2N$  parameters<sup>19</sup> that need to be estimated in the 3-dimensional model, and  $24 + 2N$  for the 2-dimensional model. Thus the estimation of the model parameters is a computationally intensive exercise. To contain the issues arising from the high dimensionality of the proposed models and associated with the stability of the large number of parameter estimates, we subdivide the estimation procedure into three steps as follows.

1. Firstly, we estimate the parameters of the interest rate process by using US Treasury yields.
2. Secondly, we estimate, separately, the parameters of the stochastic volatility model used to describe the dynamics of the commodity futures price (without a correlation structure)<sup>20</sup> by using prices of commodity derivatives.
3. Thirdly, we perform a sensitivity analysis to assess the contribution of the model correlation structure on commodity derivative prices, and we estimate the associated correlation parameters.

We estimate the parameters of the interest rate dynamics by using Kalman filter and maximum likelihood estimates. The estimation method for the futures price volatility pro-

---

<sup>19</sup>We use  $n = 3$  of each  $\xi_0, \xi, \eta, \kappa, \gamma, \rho^{xr}, \rho^{x\sigma}, \rho^{r\sigma}, \varpi, \varpi^\sigma$ ,  $N$  of each  $\lambda, \theta$  and one of each  $\bar{r}, f_0, \sigma_f, \sigma_o$ . The  $\varpi_i, \varpi_i^\sigma, f_0, \sigma_f, \sigma_o$  are the market price of risk, market price of volatility risk, initial time-homogenous futures curve, measurement noise of the futures and options, respectively.

<sup>20</sup>The correlation between futures prices and its volatility is part of the stochastic volatility model, thus it is estimated in Step 2. We assume though that the correlation  $\rho^{r\sigma}$  between interest rate process and volatility process, as well as, the correlation  $\rho^{xr}$  between interest rate process and futures price process, are both zero.

cesses  $\sigma_i^F$  and the volatility processes  $\sigma_i(t)$  involves using the extended Kalman filter and maximum likelihood, where the model is re-expressed in a state-space form which consists of the system equations and the observation equation, and then the maximum likelihood method is employed to estimate the state variables.

The system equations describe the evolution of the underlying state variables. In this model, the state vector is  $\Xi_t = \{\Xi_t^i, i = 1, 2, \dots, n\}$ , where  $\Xi_t^i$  consists of seven state variables  $x_i(t), y_i(t), z_i(t), \phi_i(t), \psi_i(t), \sigma^i(t)$  and  $\nu^i(t)$ , see (7). This system can be put in a state-space discrete evolution form as follows:

$$\Xi_{t+1} = \Phi_0 + \Phi_X \Xi_t + \omega_{t+1}. \quad (17)$$

The  $\omega_t$  with  $t = 0, 1, 2, \dots$  are independent of each other, with zero mean and with the covariance matrices conditional on time  $t$  a deterministic function of the state variable  $\sigma(t)$ . The observation equation links the vector  $\mathbf{Z}_t$  of the observable prices of futures and futures options, as described and selected in Section 3.1.2, with the vector  $\Xi_t$  of the model state variables, namely:

$$\mathbf{Z}_t = h(\Xi_t, \mathbf{u}_t), \mathbf{u}_t \sim \text{i.i.d. } N(\mathbf{1}, \Omega), \quad (18)$$

where  $\mathbf{u}_t$  is a vector of i.i.d. multiplicative Gaussian measurement errors with covariance matrix  $\Omega$ . The observation equation can be constructed by (6), which relates the logarithm of the futures prices linearly to the state variables  $x_i(t), y_i(t), z_i(t), \phi_i(t)$  and  $\psi_i(t)$ . However, equation (12) relates option prices to the state variables through nonlinear expressions. The application of the extended Kalman filter for parameter estimation involves linearising the observation function  $h$  and making the assumption that the disturbance terms  $\{\omega_t\}_{t=0,1,\dots}$  in equation (17) follow multivariate normal distributions. From the Kalman filter recursions, we can compute the likelihood function.

When interest rates are assumed to be stochastic, the interest rates needed for discounting at each future date are calculated from the estimated values of the state variables of the interest rate model. In particular, we have the estimated parameters of the interest rate model  $\bar{r}, \lambda_j, \theta_j$ , where  $j = 1, 2, 3$  (for the three-dimensional model). On each date, the



Kalman filter updates  $r_j(t)$ . With known parameters and state variables, we use equation (12) to calculate the theoretical option prices. Note that  $r_2$  and  $r_3$  are independent, so they do not appear in the main option pricing function  $G_{a,b}(y)$ . When interest rates are assumed to be deterministic, the rates used for discounting at each future date are specified by equation (15).

### *3.3. Computational details*

The program is written in Matlab. The log-likelihood function is maximised using Matlab’s “fminsearch” routine to search for the minimal point of the negative of the log-likelihood function. “fminsearch” is an unconstrained nonlinear optimisation routine and it is derivative-free. The Ricatti ordinary differential equations of the characteristic function  $\phi$  in equation (9) are solved by Matlab’s “ode23”, which is an automatic step-size Runge-Kutta-Fehlberg integration method. This method uses lower order formulae compared to other ODE methods, which can be less accurate, but the advantage is that this method is fast comparing to other methods. The integral in (12) is computed by the Gauss-Legendre quadrature formula with 19 integration points and truncating the integral at 33. We find that these numbers of integration points and truncation of the integral provide a good trade-off between computational time and accuracy. Appendix A discusses in detail the methods employed to make the estimation procedure more time efficient.

## **4. Estimation results and empirical analysis**

As part of the estimation applications, we conduct several investigations into the pricing performance of the models developed in Section 2. In the estimation part, we discuss the statistical significance and the economic significance of the estimated parameters over three distinct sample periods, characterised by different market conditions; typically low against high interest rate volatility environments. In the empirical analysis, we firstly evaluate the ability of stochastic interest rate models to improve pricing performance on long-dated crude oil derivatives compared to models with deterministic interest rate specifications. Secondly, we assess both in-sample and out-of-sample pricing performance on long-dated crude oil options with maturities up to five years. Thirdly, given the multi-dimensional nature of the

models, we investigate the sufficient number of dimensions required for the models to provide satisfactory levels of pricing performance and we discuss the trade-off between computational effort and numerical accuracy.

#### *4.1. Estimation*

The results of the three-step estimation of the proposed models, as described in Section 3.2, are discussed next.

##### *4.1.1. Interest rate process*

The estimation results of the multi-dimensional affine term structure models for the interest rate process are summarised in Table 2. We estimated the interest rate process for  $N = 1, 2, 3$ , to determine the number of model dimensions required to provide a satisfactory fit to interest rate data. The results reveal that the three-dimensional affine term structure model provides the best fit for all maturities, as RMSE (root-mean-square-errors) are consistently lower for all maturities. This is consistent with results by Litterman and Scheinkman (1991) and Fan, Gupta, and Ritchken (2007) in the swaption market. Therefore, we consider the three-dimensional version of interest rate models in the subsequent analysis.

The estimated long-term mean level  $\bar{r}$  in the first period is 4.96%, which is much higher than the long-term mean level of 0.9519% estimated in the second period and of  $-0.20\%$  estimated in the third period. These parameter estimates are consistent with the statistical properties of the interest rates over these three periods, see Table 1. Similarly, the estimated volatility parameters  $\theta_i$  of the interest rate process reflect the observed variation in the interest rates markets over the three periods. Specifically, the August 2005–July 2007 period is characterised by high levels of interest rates and relatively high volatility (ranging from 0.8%-1.5%), the July 2007–July 2009 period is characterised by lower interest rates with considerable volatility (ranging from 1.28%-2.77%), while the January 2011–December 2012 period exhibits much lower interest rates and very low volatility.

##### *4.1.2. Futures price process*

Next, we estimate the stochastic volatility models developed in Section 2, by assuming zero correlation between the futures price process and the interest rate process and between

	Period 1: Aug 05–Jul 07					
	N=1	N=2		N=3		
	$i = 1$	$i = 1$	$i = 2$	$i = 1$	$i = 2$	$i = 3$
$\lambda_{0i}$	0.0492 (0.0005)	-0.0401 (0.0002)	3.9081 (0.5768)	0.0752 (0.0047)	0.2766 (0.0007)	1.5860 (0.0043)
$\theta_i$	0.0058 (0.0011)	0.0058 (0.0004)	0.0164 (0.0020)	0.0152 (0.0008)	0.0080 (0.0004)	0.0130 (0.0012)
$\bar{r}$	4.1592% (0.0025)	3.9112% (0.0016)		4.9611% (0.0010)		
log L	11255	12803		13240		
RMSE 1yr	4.1033%	1.0425%		0.4885%		
RMSE 2yr	1.1798%	0.4110%		0.2823%		
RMSE 3yr	0.6045%	0.3392%		0.1559%		
RMSE 5yr	0.5367%	0.1346%		0.0476%		

	Period 2: Jul 07–Jun 09					
	N=1	N=2		N=3		
	$i = 1$	$i = 1$	$i = 2$	$i = 1$	$i = 2$	$i = 3$
$\lambda_{0i}$	0.1525 (0.0034)	-0.2616 (0.0216)	0.3158 (0.0237)	-0.1572 (0.0065)	0.4283 (0.0021)	0.1247 (0.0010)
$\theta_i$	0.0183 (0.0015)	0.0079 (0.0014)	0.0188 (0.0020)	-0.0277 (0.0002)	0.0128 (0.0010)	-0.0235 (0.0001)
$\bar{r}$	1.7960% (0.0010)	1.9715% (0.0020)		0.9519% (0.0003)		
log L	11344	12024		13735		
RMSE 1yr	41.7333%	16.2820%		11.9594%		
RMSE 2yr	16.0380%	3.6141%		1.8584%		
RMSE 3yr	6.5301%	3.1575%		1.3201%		
RMSE 5yr	4.7623%	1.1613%		0.1860%		

	Period 3: Jan 11–Dec 12					
	N=1	N=2		N=3		
	$i = 1$	$i = 1$	$i = 2$	$i = 1$	$i = 2$	$i = 3$
$\lambda_{0i}$	-0.6491 (0.0038)	-0.3913 (0.0093)	0.5144 (0.0585)	0.1223 (0.0116)	-0.4987 (0.0034)	-1.2297 (0.0303)
$\theta_i$	0.0010 (0.0001)	0.0027 (0.0001)	0.0025 (0.0003)	0.0021 (0.0001)	0.0019 (0.0003)	3.98E-05 (0.0000)
$\bar{r}$	0.0700% (0.0000)	-0.2828% (0.0001)		-0.2018% (0.0004)		
log L	11602	13057		13595		
RMSE 1yr	47.6749%	22.5815%		15.3528%		
RMSE 2yr	9.8461%	5.6910%		3.8036%		
RMSE 3yr	4.7325%	2.8091%		1.2320%		
RMSE 5yr	0.9182%	0.3006%		0.0661%		

Table 2: **Parameter estimates of the interest rate process.** The table displays the maximum-likelihood parameter estimates (with the standard errors in parenthesis) and the root-mean-square errors (RMSE) of multi-dimensional affine term structure models for  $N = 1, 2, 3$  over three periods, August 2005 – July 2007, July 2007–June 2009 and January 2011 – December 2012.

the interest rate process and the stochastic volatility process (i.e.  $\rho_i^{xr} = \rho^{r\sigma} = 0$ ) over the three periods. The estimation results of fitting the two-dimensional model and the three-dimensional model, to crude oil futures and options are shown in Table 3 and Table 4, respectively. Thus, the tables display the estimates obtained under the assumption of  $\rho_i^{xr} = \rho^{r\sigma} = 0$ . The correlation coefficients are estimated separately as discussed in Section 4.1.4.

The estimated parameter values show that the volatility factors driving the futures price process are highly persistent (low  $\kappa_i$ ) and they are typically exponentially decaying (low  $\xi_i$  estimates) or may exhibit humps, when  $\xi_i$  estimates are large, as it has been empirically shown also in Chiarella et al. (2013). The high values of the correlation coefficients  $\rho^{x\sigma}$  between innovations in the futures price process and the volatility process reflect the existence of unspanned volatility components in the crude oil futures markets, see Trolle and Schwartz (2009). Factors with higher  $\eta_i$  drive mainly the short end of the term structure. Furthermore, in line with Dennis, Mayhew, and Stivers (2006) and Chiarella et al. (2016), the negative correlations in Period 1 and 2 imply a negative return–volatility relationship in the crude oil futures markets over these periods leading up to the GFC, where systematic market wide shocks impact the market. The more quiet Period 3 displays a positive return–volatility relationship consistent with the fact that the market was mostly in backwardation.<sup>21</sup> The market price of volatility risk  $\varpi_i^\sigma$  in the crude oil market has negative components, consistent with several empirical studies, see for instance Doran and Ronn (2008) and Chiang, Hughen, and Sagi (2015).

#### 4.1.3. Sensitivity analysis of the correlations

A sensitivity analysis of the correlation  $\rho^{xr}$  between the interest rate process and the futures price process, as well as the correlation  $\rho^{r\sigma}$  between the interest rate process and the stochastic volatility process, is performed to determine their impact on futures and option

---

<sup>21</sup>According to Dennis et al. (2006) and Chiarella et al. (2016), commodity futures markets with low volatility are driven predominantly by commodity specific effects, for example, crude oil behaves as a consumption asset (inverted futures markets) and gold as an investment asset, postulating respectively, positive and negative return–volatility relationships (i.e., positive and negative correlation). However, in high volatility market conditions, commodity markets are affected also by market-wide shocks and may behave differently. Thus, over the Period 1 and 2, where the GFC effects were dominating, the crude oil markets experienced extended periods of contango market conditions (meaning, futures prices exceed spot prices), which is consistent with negative correlation.

prices. By using the estimated model parameters of the zero-correlation models, see Table 3 and Table 4, we price call options with a different correlation structure. In our analysis, we select an at-the-money (ATM) option with a maturity of 4,000 days and a strike of \$100. These numbers are chosen for illustration purposes only and we present the results for two of the periods. The results are presented in Table 5.

From this analysis, we draw several conclusions. Firstly, the impact of the correlation  $\rho^{r\sigma}$  between the interest rate process and the stochastic volatility process on the option prices is insignificant even for maturity as long as 4,000 days (see Table 5). For  $\rho^{xr} = 0$ ,<sup>22</sup> we see that the percentage difference of the option price (comparing  $\rho^{r\sigma} = 0.50$  and  $\rho^{r\sigma} = 0$ ) is  $(10.51 - 10.416)/10.51 = 0.894\%$  for the period 2005 – 2007 and the percentage difference of the option price (comparing  $\rho^{r\sigma} = 0.50$  and  $\rho^{r\sigma} = 0$ ) is  $(28.201 - 28.187)/28.201 = 0.05\%$  for the period 2011 – 2012.

Secondly, the correlation  $\rho^{xr}$  between the interest rate process and the underlying futures price process has a noticeable impact on the option prices. More specifically, due to the differences in the interest rate volatilities estimated for these periods, the impact of the correlation on the option prices in the period 2005 – 2007 is more than twice as large as in the period from 2011 – 2012. For instance, the percentage difference of the option price (comparing  $\rho^{xr} = 0$  and  $\rho^{xr} = -0.50$ ) is  $(10.51 - 10.857)/10.51 = -3.30\%$  for the period 2005 – 2007 and the percentage difference of the option price (comparing  $\rho^{xr} = 0$  and  $\rho^{xr} = -0.50$ ) is  $(28.201 - 28.591)/28.201 = -1.38\%$  for the period 2011 – 2012. Consequently, as the impact of the correlation between the stochastic interest rate process and the stochastic volatility to the option price is negligible, we set  $\rho_1^{r\sigma} = \rho_2^{r\sigma} = \rho_3^{r\sigma} = 0$  and we only estimate  $\rho_i^{xr}$ .

#### 4.1.4. The correlation $\rho_i^{xr}$

Table 6 and Table 7 present the estimates of the correlation coefficient  $\rho_i^{xr}$  for the two-dimensional and the three-dimensional model, respectively. The remaining parameter es-

---

<sup>22</sup>Throughout this section, we use the shorthand notation  $\rho^{xr} = 0$  to refer to  $\rho_1^{xr} = \rho_2^{xr} = \rho_3^{xr} = 0$ , and the same interpretation for  $\rho^{r\sigma}$ .

timates are the same as in Table 3 and Table 4. The correlation coefficients between the interest rate process and the futures price process reach quite high values in absolute terms, ranging from  $-0.69$  to  $0.54$ , for the two-dimensional model and from  $-0.62$  to  $0.59$ , for the three-dimensional model, underscoring the important (though not invariant) relationship between interest rates and futures prices. We observe that, especially over the high interest rate volatility periods of 2005 – 2007 and 2007 – 2009, these correlations are always negative. This is consistent with studies such as Akram (2009), Arora and Tanner (2013) and Frankel (2014), which provide empirical evidence for a negative relationship between oil prices and interest rates. Akram (2009) conducts an empirical analysis based on structural VAR models estimated on quarterly data over the period 1990–2007. One of his results suggests that there is a negative relationship between the real oil prices and real interest rates. Furthermore, Arora and Tanner (2013) suggest that oil prices consistently fall with unexpected rises in short-term real interest rates through the whole sample period from 1975 to 2012. They also found that oil prices have become more responsive to long-term real interest rates over time. Frankel (2014) presents and estimates a “carry trade” model of crude oil prices and other storable commodities. Their empirical results support the hypothesis that low interest rates contribute to the upward pressure on real commodity prices via a high demand for inventories. Even though our empirical analysis does not refer to the correlations between the actual financial observables as the above econometric studies do, it reveals a negative correlation between innovations of the crude oil futures prices and innovations in the interest rate process. This implies that crude oil futures prices and crude oil spot prices have a similar response to changes in the interest rates.

#### *4.2. Pricing performance on long-dated crude oil derivatives*

The in-sample and out-of-sample pricing performance of stochastic interest rate models on long-dated crude oil derivatives is discussed next, by comparing to models with deterministic interest rate specifications.

##### *4.2.1. Model Fit*

Comparing Table 3 and Table 4, we note that overall the three-dimensional models outperform the two-dimensional models, by consistently providing better log-likelihood (log

L) and Akaike’s information criterion (AIC) values over all three sample periods. More specifically, the three-dimensional models provide a better fit to futures prices (compared to the two-dimensional models) as they produce lower RMSE over all three sample periods. However, when fitting to the implied volatility, the improvement in the RMSE between the two-factor models and the three-factor models is marginal, in the pre-GFC Period 1 and the post-GFC Period 3. For the period 2005 – 2007, the two-dimensional model seems to fit slightly better the very long maturities of the 4-year and 5-year implied volatility. As it has been also shown in Schwartz and Smith (2000) and Cortazar et al. (2016a), increasing the dimensionality of multi-dimensional models does not always improve pricing on long-dated commodity contracts. Yet, the fit to implied volatilities over Period 2 (covering the GFC period, where the crude oil market experienced extreme variations) is improving considerably between two-dimensional and three-dimensional models i.e., by 104 basis points (from 3.9361% to 2.8921%) for 4-year contracts and by 115 basis points (from 3.8580% to 2.7094%) for 5-year contracts. Thus, when markets experience significant volatility, three-dimensional futures price models tend to provide a better fit to implied volatility of long-dated contracts, while under typical market conditions, two-dimensional futures price models may suffice.

Furthermore, we compare the model fit of the models ignoring correlation between the futures price process and interest rate process, thus assuming  $\rho_i^{xr} = 0$ ,<sup>23</sup> see Table 3 and Table 4, to the models that take the correlation coefficients into account, as presented in Table 6 and Table 7, respectively. The models which allow for correlation provide both better log L and AIC, underscoring the importance of the correlations to improving model fit. In Period 3, the inclusion of the correlation  $\rho_i^{xr}$  marginally improves the fit of both futures prices and implied volatility. This is mainly because during that period interest rates and their volatility were very low and therefore the interest rate process has very little impact on the option prices. However, we observe a substantial improvement in the RMSE of the implied volatility in Period 2 (GFC period), when incorporating the correlation  $\rho_i^{xr}$ . For instance, in the three-dimensional model, the RMSE of the implied volatility of options with five years to maturity improves by 52 basis points (from 2.7094% to 2.1850%) for Period 2 and by

---

<sup>23</sup>Recall that according to the sensitivity analysis in Section 4.1.3, we also assume that the correlation coefficient  $\rho_i^{r\sigma} = 0$ .

26 basis points for Period 1. We also observe a moderate improvement over the pre-GFC period (2005–2007) to the fit of implied volatilities of all maturities and both models. Thus, in periods of high uncertainty, including more factors and taking correlations into account improve the fit to long-dated crude oil derivatives. Over periods of typical market conditions, adding more factors does not bring improvement, but including correlations does.

#### *4.2.2. Deterministic vs stochastic interest rates*

Next, we examine the impact of including stochastic interest rates when pricing long-dated crude oil derivatives. It is well documented that stochastic volatility alone improves pricing performance on long-dated equity derivatives, see Bakshi et al. (2000), as well as long-dated commodity derivatives, see Cortazar et al. (2016a). Yet, stochastic interest rates are important when considering long-dated commodity commitments, see Hilliard and Reis (1998) and Grzelak et al. (2012). Thus, we compare the pricing performance of the proposed three-factor futures price model that incorporates both stochastic volatility and stochastic interest rates with corresponding models that assume deterministic interest rates (i.e. discounted by the corresponding Treasury yields).

We also use the model parameters estimated in the previous section to assess out-of-sample performance by re-running the scenarios with extended data. The data extends from 1<sup>st</sup> August 2007 to 31<sup>st</sup> December 2007 for the first period, from 1<sup>st</sup> July 2009 to 31<sup>st</sup> December 2009 for the second period and from 1<sup>st</sup> January 2013 to 30<sup>th</sup> June 2013 for the third period. Figure 6 displays the average of the RMSE across all maturities between the volatility implied by the market option prices on that day and the implied volatility from the estimated model (as described in Section 3.1.2) for the three sample periods used in our analysis. Figure 7 and Figure 8 display the average of the daily RMSE between the implied volatility from the estimated models and the observed implied volatility for maturities of 2, 3 and 4, 5 years, respectively. Since the liquidity of the crude oil long-dated contracts is concentrated in the December contracts, we assess the model fit to December contracts which are in the second, third, fourth and fifth year to maturity. The results of both the in-sample and out-of-sample analysis are summarised in Table 8.

We observe that models that incorporate stochastic interest rates typically improve pric-



ing performance, yet the improvement becomes more evident as the maturity of the option contracts increases. For maturities of up to two years, the improvement is minimal or not discernible for both the in-sample and the out-of-sample analysis. Indeed, Figure 7 and Figure 8 display occasions, where the deterministic interest rate specifications lead to lower (in-sample) RMSE over part of the sample compared to the stochastic interest rate specifications.<sup>24</sup> Yet, stochastic interest rate specifications have a clear advantage over periods where markets experience extreme variation, see the middle panels of Figure 7 and Figure 8. Table 8 also reveals that the in-sample improvement on RMSE of implied volatilities reaches 100 basis points, on average, for maturities up to 5 years in the GFC period, and 47 basis points in the pre-GFC period. In contrast, in Period 3, which is characterised by extremely low interest rate volatilities, stochastic interest rate models do not provide any improvement in the in-sample pricing performance but improve out-of-sample pricing performance, especially for long-dated contracts. Most specifically, the out-of-sample improvement on RMSE of implied volatilities reaches 80 basis points, on average, for maturities up to 5 years in the pre-GFC period and 35 in the post-GFC period (Period 3).

Thus, stochastic interest rates become relevant and important to pricing long-dated contracts, when the interest rate volatility is high. During the period January 2011–December 2012, where the interest rate volatility was low, marginal improvement in pricing performance is observed for all maturities. Thus, when interest rate are not volatile, stochastic interest rates will not improve in-sample pricing performance, but they do contribute to improving out-of-sample pricing performance to some extent.

#### *4.2.3. Robustness of the results*

To assess the robustness of the estimation and of the results, several measures have been considered. The model performance is assessed through log-likelihood and the Akaike criterion, and both lead to the same conclusions. Also, all estimates are statistically significant,

---

<sup>24</sup>By construction, the stochastic interest rate specification will always lead to a better fit in-sample when the entire sample is considered, but since the models are estimated using the full sample, this does not necessarily preclude better RMSE for the simpler model on particular sub-samples. The key test, however, remains the out-of-sample performance.

as is evident from Table 3 and Table 4.

Secondly, the three-step estimation approach has been employed to reduce the dimensionality of the estimation problems and improve the stability of the parameter estimates.<sup>25</sup> The stochastic interest rate process has been estimated separately to the commodity derivatives price model, as the former involves fitting to interest rate data only. The estimated models retain the market observed features at the corresponding sample periods, such as high mean interest rate in the pre-GFC period, low interest rate volatility in the post-GFC period and high interest rate volatility during the GFC. In addition, the correlation coefficient that links these two markets has been estimated in a third estimation step. This lowers the computation cost and allows us to concentrate on the impact of the correlation. The impact of the correlation coefficient on the proposed models has been subjected to a sensitivity analysis, see Section 4.1.3. Table 3 and Table 6, as well as, Table 4 and Table 7 present the estimation results of a restricted model of zero correlation and a model with non-zero correlation, respectively. They reveal the manner in which a model with non-zero correlation coefficient can provide a better fit to market data. There is also consistency of the estimated model coefficients with the descriptive statistics, i.e. the empirically observed negative correlation between innovations in the interest rates and futures prices, a result that holds for all three sample periods.

Thirdly, the optimisation procedure involved in the estimation was repeated with numerous different plausible initial parameter guesses to reduce the risk of not reaching the global optimum, in a manner similar to Trolle and Schwartz (2009) and Chiarella et al. (2013).

Fourthly, the results are economically significant. Taking into consideration that the current bid-ask spreads for 0.5-year, 1.5-year and 2.5-year options, expressed in terms of implied volatility differences, are around 1.06%, 1.41% and 2.67%<sup>26</sup> respectively, the improvement

---

<sup>25</sup>Conceivably, further gains could be achieved by combining steps into a larger, simultaneous estimation. However, this is at the risk of introducing instability, for marginal potential gain, so we did not pursue this avenue of inquiry further.

<sup>26</sup>These numbers represent the average bid-ask spreads (expressed in terms of implied volatilities differences) of the corresponding maturities of Dec-2016, Dec-2017 and Dec-2018 option contracts observed in June 2016. For example, for the implied volatility of the Dec-2017 contracts, we use the Dec-2017 futures prices, 1.5-year Treasury yields and a minimisation routine to match the implied volatility to the Dec-2017 futures option prices.

in the maximum absolute differences between the volatility implied by the market crude oil option prices and the model implied volatility (of the deterministic interest rate model and the stochastic interest rate model) observed in the in-sample period and the out-of-sample period, are of similar magnitude as the implied volatility differences of the bid-ask spread prices, especially in periods of high interest rate volatility such as the pre-GFC and GFC periods, see Table 9.<sup>27</sup> Furthermore, over the GFC, the stochastic interest rate model provides lower absolute model/market differences compared to the deterministic interest rate model, in particular out of sample and for longer maturities. These observations underscore the importance of stochastic interest rate models for pricing long-dated commodity derivatives. In particular, given that when hedging long-dated commodity derivative contracts this effect would accumulate over the repeated re-balancing of the dynamic hedging strategy, one would expect the inclusion of stochastic interest rates in the model to improve its hedging performance.

The use of the Kalman filter estimation method (standard or extended) builds on extensive literature on this topic, in which it has been widely and successfully applied, including in the context of models similar to the one at hand.<sup>28</sup> It is recognised that the method involves simplifying assumptions (such as linearity of the observation equation, Gaussian state space, etc.), which may lead to a suboptimal estimation of the model parameters. However, this does not detract from our results, most importantly the fact that the more advanced stochastic volatility–stochastic interest rate model outperforms the corresponding

---

<sup>27</sup>This compares favourably to other studies. For example, Chiang et al. (2015) report RMSE which is about ten times higher than the current bid–ask spread.

<sup>28</sup>The Kalman filter is an estimation method which has been extensively used in the literature of estimating model parameters from market data, including commodity futures and options. The quasi-maximum likelihood is used for the estimation (calibration) of the parameter values. This assumes a Gaussian distribution to approximate the true distribution of the pricing errors, see Harvey (1989). For applications of the Kalman filter to commodity markets, where estimation of model parameters by fitting to futures prices typically implies a linear measurement equation, see Schwartz (1997), Schwartz and Smith (2000), Dempster et al. (2008), Dempster, Medova, and Tang (2012), and Casassus, Liu, and Tang (2013). Model parameter estimation by fitting to futures and options data typically requires an extended version of the classical Kalman filter, where the measurement equation for fitting to option prices is linearized, for instance see Trolle and Schwartz (2009), Chiarella et al. (2013) and Cortazar, Gutierrez, and Ortega (2016b). Duffee and Stanton (2012a) have demonstrated that for models with dynamic affine term structures (see Duffee and Kan (1996)), the Kalman filter (and the extended Kalman filter) outperforms other standard estimation approaches such as Maximum Likelihood and Efficient Method of Moments.

deterministic interest rate model, in particular out of sample, during periods of high interest rate volatility, and for longer maturities. This is despite any shortcomings of the estimation method, and a method which overcomes these shortcomings might strengthen the results. However, given that the Kalman filter remains the method of choice in the literature, we leave this point to further research.

## 5. Conclusion

We make both theoretical and empirical contributions to pricing long-dated commodity derivatives. Our theoretical contributions lie in developing futures price models for commodity derivatives that allow for both stochastic volatility and stochastic interest rates, as well as a correlation structure between the underlying variables. The proposed models have several advantages. Firstly, they are multi-dimensional futures price models that model directly the entire term structure of the futures prices and avoid the need to model the unobserved convenience yield directly. Secondly, the models are consistent with empirically observed market features such as unspanned volatility and a variety of volatility structures. Thirdly, the models are tractable, as they provide affine structures for futures prices and quasi analytical solutions for commodity options. Thus the models can be used for estimation and calibration purposes. Fourthly, the models are well suited for pricing and hedging applications involving long-dated commodity derivatives, as they allow for both stochastic volatility and stochastic interest rates and can fit the term structure of derivative prices.

Our empirical contributions include a study on the ability of stochastic volatility–stochastic interest models to improve pricing of long-dated crude oil derivatives. We estimate the model parameters from historical time series of both crude oil futures prices and option prices, over three periods characterised by different interest rate market conditions (in terms of interest rate volatility). To this end, we further conduct a sensitivity analysis on the contribution of the model parameters such as the correlations and the interest rate volatility to commodity derivative prices. We also evaluate the improvement of pricing performance on long-dated crude oil derivatives achieved, by including stochastic interest rate dynamics (beyond stochastic volatility) in the model. Additionally, we investigate the number of dimensions required

for the models to provide satisfactory levels of pricing performance. The empirical analysis leads to the following conclusions. First, stochastic volatility futures price models that incorporate stochastic interest rates typically improve pricing performance on long-dated crude oil derivatives (compared to models with deterministic interest rate specifications), an improvement that is especially evident over periods of high interest rate volatility and for longer-maturity contracts. Second, the pricing performance improvement is more prominent and consistent in the out-of-sample fitting. Third, in periods of high uncertainty, including more factors and taking correlations into account improve the fit to long-dated crude oil derivatives. Over periods of typical market conditions, adding more factors does not bring improvement, but it is important to consider the correlations. Fourth, there is empirical evidence for a negative correlation between the futures price process and the interest rate process, especially over periods of high interest rate volatility. Fifth, the correlation between the futures price volatility process and the interest rate process has negligible impact on the pricing of long-dated crude oil contracts.

These empirical results may well provide useful insights to practitioners. Increasing the dimension of multi-dimensional futures price commodity models may improve the model fit to historical market data, but it comes with additional computational effort. Our results show that for pricing long-dated crude oil commodities, two-dimensional stochastic volatility-stochastic interest rate models are adequate under typical market conditions but they do not suffice under market turmoil. The key point, however, is that in periods of high interest rate volatility (at levels well in line with recent historical precedent), modelling stochastic interest rates and including correlations do matter for pricing longer maturity crude oil derivatives. Otherwise, using deterministic interest rates will be sufficient, but it is unwarranted to remain complacent in assuming that interest rate volatilities will remain too low to matter.

## **Acknowledgements**

We thank seminar participants at the 2014 Quantitative Methods in Finance conference, the 2016 Energy and Commodity Finance conference and the UTS Finance Discipline Group internal research seminars for fruitful discussions and helpful suggestions. We would also wish to thank the late Carl Chiarella, Boda Kang, Antonios Papapantoleon, two anonymous

reviewers and the Australian Research Council for financial support with purchasing the data (DP 1095177, The Modelling and Estimation of Volatility in Energy Markets).

## Appendix A. Methods to reduce computational time

One of the biggest challenges is the formidable amount of computational time required for the estimation of parameters. Although this model admits quasi-analytic solutions for option pricing, complex-valued numerical ODE approximation together with complex-valued numerical integration are needed for each option price. Furthermore, for each day of the crude oil data, a numerical Jacobian needs to be calculated for the linearisation of option prices in the Kalman filter update. To complete one day of the data, which typically involves the calculation of around 70 options and its Jacobian for the Kalman filter update, may take 10 to 15 minutes on a desktop running a second generation quad-core i7 processor. So that is about 5,000 to 7,500 minutes for the program to process two years (about 500 trading days) of data in order to calculate one log-likelihood. Matlab’s “fminsearch” routine may take several hundreds of iterations for it to converge to a local maximum.

The key observation to massively reduce the computational time is that, given a set of parameters of our model, the characteristic function  $\phi(t; a + \mathbf{i}bu, T_o, T)$  is a function of  $a, b, u, T_o, T$ .  $T_o$  and  $T$  are set to be the same because all the crude oil options traded in CME expire only a few days before the underlying futures contracts. So for each iteration we precalculate six tables for the characteristic functions. These are  $\phi(t; 0, T, T)$ ,  $\phi(t; 1, T, T)$ ,  $\phi(t; 1 - \mathbf{i}u, T, T)$ ,  $\phi(t; -\mathbf{i}u, T, T)$ ,  $\phi(t; 1 + \mathbf{i}u, T, T)$  and  $\phi(t; \mathbf{i}u, T, T)$  where values of the variable  $u$  are determined by the 19 integration points on the interval from 0 to 33 calculated using Gauss-Legendre quadrature, and the values of the maturity  $T$  are 14, 15, 16,  $\dots$ , 1850, because the shortest maturity is only 14 days and the longest maturity is five years. Observing the fact that each calculation of the characteristic function is independent of the others, we also take advantage of the parallel toolbox available in Matlab by using Matlab’s “parfor” loop. Applying all of the above reduces the time required to process one iteration (2 years of data) from 5,000 to 7,500 minutes to around 2 minutes. The total time required for the parameters to converge to an optimum can still take a few days.

## References

- Akram, Q. F., 2009. Commodity prices, interest rates and the dollar. *Energy Economics* 31 (6), 838–851.
- Amin, K. I., Jarrow, R. A., 1992. Pricing options on risky assets in a stochastic interest rate economy. *Mathematical Finance* 2 (4), 217–237.
- Amin, K. I., Ng, V., 1993. Option valuation with systematic stochastic volatility. *Journal of Finance* 48, 881–910.
- Arora, V., Tanner, M., 2013. Do oil prices respond to real interest rates? *Energy Economics* 36, 546–555.
- Bakshi, G., Cao, C., Chen, Z., 1997. Empirical performance of alternative option pricing models. *The Journal of Finance* 52 (5), 2003–2049.
- Bakshi, G., Cao, C., Chen, Z., 2000. Pricing and hedging long-term options. *Journal of Econometrics* 94, 277–318.
- Ballotta, L., Haberman, S., 2003. Valuation of guaranteed annuity conversion options. *Insurance: Mathematics and Economics* 33 (0), 87–108.
- Barone-Adesi, G., Whaley, R. E., 1987. Efficient analytic approximation of American option values. *The Journal of Finance* 42 (2), 301–320.
- Björk, T., Blix, M., Landén, C., 2006. On finite dimensional realizations for the term structure of futures prices. *International Journal of Theoretical and Applied Finance* 9 (3), 281–314.
- Björk, T., Landén, C., Svensson, L., 2004. Finite-dimensional Markovian realizations for stochastic volatility forward-rate models. *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences* 460 (2041), 53–83.
- Black, F., Scholes, M., 1973. The pricing of options and corporate liabilities. *The Journal of Political Economy* 81 (3), 637–654.
- Casassus, J., Collin-Dufresne, P., 2005. Stochastic convenience yield implied from commodity futures and interest rates. *The Journal of Finance* 60 (5), 2283–2331.
- Casassus, J., Liu, P., Tang, K., 2013. Economic linkages, relative scarcity, and commodity futures returns. *The Review of Financial Studies* 26, 1324–1362.
- Chen, R.-R., Scott, L., 1993. Maximum likelihood estimation for a multifactor equilibrium model of the term structure of interest rates. *The Journal of Fixed Income* 3 (3), 14–31.
- Chiang, I., Hughen, W. K., Sagi, J. S., 2015. Estimating oil risk factors using information from equity and derivatives markets. *The Journal of Finance* 70, 769–804.

- Chiarella, C., Kang, B., Nikitopoulos, C. S., Tô, T.-D., 2013. Humps in the volatility structure of the crude oil futures market: New evidence. *Energy Economics* 40, 989–1000.
- Chiarella, C., Kang, B., Nikitopoulos, C. S., Tô, T.-D., 2016. The return-volatility relation in commodity futures markets. *Journal of Futures Markets* 36 (2), 127–152.
- Chiarella, C., Kwon, O. K., 2003. Finite dimensional affine realisations of HJM models in terms of forward rates and yields. *Review of Derivatives Research* 6 (2), 129–155.
- Cortazar, G., Gutierrez, S., Ortega, H., 2016a. Empirical performance of commodity pricing models: When is it worthwhile to use a stochastic volatility specification? *Journal of Futures Markets* 36 (5), 457–487.
- Cortazar, G., Gutierrez, S., Ortega, H., 2016b. Empirical performance of commodity pricing models: When is it worthwhile to use a stochastic volatility specification? *Journal of Futures Markets* 36 (5), 457–487.
- Cortazar, G., Naranjo, L., 2006. An n-factor Gaussian model of oil futures prices. *Journal of Futures Markets* 26 (3), 243–268.
- Cortazar, G., Schwartz, E. S., 2003. Implementing a stochastic model for oil futures prices. *Energy Economics* 25 (3), 215–238.
- Cox, J., Ingersoll, J. E., Ross, S. A., 1985. A theory of the term structure of interest rates. *Econometrica* 53, 385–407.
- Cox, J. C., Ingersoll, J. E., Ross, S. A., 1981. The relation between forward prices and futures prices. *Journal of Financial Economics* 9, 321–346.
- Dempster, M., Medova, E., Tang, K., 2008. Long term spread option valuation and hedging. *Journal of Banking and Finance* 32 (12), 2530–2540.
- Dempster, M. A. H., Medova, E., Tang, K., 2012. Determinants of oil futures prices and convenience yields. *Quantitative Finance* 12, 1795–1809.
- Dennis, P., Mayhew, S., Stivers, C., 2006. Stock returns, implied volatility innovations, and the asymmetric volatility phenomenon. *Journal of Financial and Quantitative Analysis* 41 (2), 381–406.
- Doran, J. S., Ronn, E. I., 2008. Computing the market price of volatility risk in the energy commodity market. *Journal of Banking and Finance* 32, 2541–2552.
- Duffee, G. R., Stanton, R. H., 2012a. Estimation of dynamic term structure models. *Quarterly Journal of Finance* 2 (2), 1250008 (51 pages).
- Duffee, G. R., Stanton, R. H., 2012b. Estimation of dynamic term structure models. *The Quarterly Journal of Finance* 2 (2), 1250008.
- Duffie, D., Kan, R., 1996. A yield-factor model of interest rates. *Mathematical Finance* 6 (4), 379–406.



- Duffie, D., Pan, J., Singleton, K., 2000. Transform analysis and asset pricing for affine jump-diffusions. *Econometrica* 68 (6), 1343–1376.
- Fan, R., Gupta, A., Ritchken, P., 2007. On pricing and hedging in the swaptions market: How many factors really? *Journal of Derivatives* 15 (1), 933.
- Frankel, J. A., 2014. Effects of speculation and interest rates in a “carry trade” model of commodity prices. *Journal of International Money and Finance* 42, 88–112.
- Fusai, G., Marena, M., Roncoroni, A., 2008. Analytical pricing of discretely monitored a sian-style options: Theory and applications to commodity markets. *Journal of Banking and Finance* 32 (10), 2033–2045.
- Gallant, A. R., Hsieh, D., Tauchen, G., 1997. Estimation of stochastic volatility models with diagnostics. *Journal of Econometrics* 81 (1), 159–192.
- Geman, H., 2005. *Commodities and Commodity Derivatives: Modeling and Pricing for Agricultural, Metals and Energy*. Wiley Finance, London, UK.
- Geman, H., Nguyen, V., 2005. Soybean inventory and forward curve dynamics. *Management Science* 51 (7), 1076–1091.
- Gibson, R., Schwartz, E., 1990. Stochastic convenience yield and the pricing of oil contingent claims. *Journal of Finance* 45 (3), 959–976.
- Grzelak, L. A., Oosterlee, C. W., 2011. On the Heston model with stochastic interest rates. *Journal on Financial Mathematics* 2 (1), 255–286.
- Grzelak, L. A., Oosterlee, C. W., van Weeren, S., 2012. Extension of stochastic volatility equity models with the Hull–White interest rate process. *Quantitative Finance* 12 (1), 89–105.
- Harvey, A., 1989. *Forecasting, structural time series models and the Kalman Filter*. Cambridge University Press, Cambridge, UK.
- Heath, D., Jarrow, R., Morton, A., 1992. Bond pricing and the term structure of interest rates: A new methodology for contingent claim valuations. *Econometrica* 60 (1), 77–105.
- Hilliard, J. E., Reis, J., 1998. Valuation of commodity futures and options under stochastic convenience yields, interest rates, and jump diffusions in the spot. *Journal of Financial and Quantitative Analysis* 33 (01), 61–86.
- Hull, J., White, A., 1990. Pricing interest-rate derivative securities. *The Review of Financial Studies* 3 (4), 573–592.
- Karlsson, P., Pilz, K. F., Schlögl, E., 2016. Calibrating market model to commodity and interest rate risk. Tech. rep., <http://ssrn.com/abstract=2773974>.
- Lipton, A., Sepp, A., 2008. Stochastic volatility models and Kelvin waves. *Journal of Physics A: Mathematical and Theoretical* 41, 1–23.

- Litterman, R. B., Scheinkman, J., 1991. Common factors affecting bond returns. *The Journal of Fixed Income* 1 (1), 54–61.
- Musiela, M., Rutkowski, M., 1997. *Martingale Methods in Financial Modelling*, 1st Edition. Vol. 36 of *Applications of Mathematics: Stochastic Modelling and Applied Probability*. Springer.
- Nelson, C. R., Siegel, A. F., 1987. Parsimonious modeling of yield curves. *Journal of Business* 60 (4), 473–489.
- Pilz, K. F., Schlögl, E., 2013. A hybrid commodity and interest rate market model. *Quantitative Finance* 13 (4), 543–560.
- Rindell, K., 1995. Pricing of index options when interest rates are stochastic: An empirical test. *Journal of Banking and Finance* 19 (5), 785–802.
- Schöbel, R., Zhu, J., 1999. Stochastic volatility with an Ornstein-Uhlenbeck process: An extension. *European Finance Review* 3 (1), 23–46.
- Schrager, D., Pelsler, A., 2004. Pricing rate of return guarantees in regular premium unit linked insurance annuity conversion options. *Insurance: Mathematics and Economics* 35 (0), 369–398.
- Schwartz, E., 1997. The stochastic behavior of commodity prices: Implications for valuation and hedging. *Journal of Finance* 52 (3), 923–973.
- Schwartz, E., Smith, J. E., 2000. Short-term variations and long-term dynamics in commodity prices. *Management Science* 46 (7), 893–911.
- Stein, E. M., Stein, J. C., 1991. Stock price distributions with stochastic volatility: An analytic approach. *Review of Financial Studies* 4 (4), 727–752.
- Trolle, A. B., Schwartz, E. S., 2009. Unspanned stochastic volatility and the pricing of commodity derivatives. *The Review of Financial Studies* 22 (11), 4423–4461.
- van Haastrecht, A., Lord, R., Pelsler, A., Schrager, D., 2009. Pricing long-dated insurance contracts with stochastic interest rates and stochastic volatility. *Insurance: Mathematics and Economics* 45 (3), 436–448.
- van Haastrecht, A., Pelsler, A., 2011. Accounting for stochastic interest rates, stochastic volatility and a general correlation structure in the valuation of forward starting options. *Journal of Futures Markets* 31 (2), 103–125.

	Period 1		Period 2		Period 3	
	Aug 2005–Jul 2007		Jul 2007–Jun 2009		Jan 2011–Dec 2012	
	$i = 1$	$i = 2$	$i = 1$	$i = 2$	$i = 1$	$i = 2$
$\xi_{0i}$	0.0564 (0.0002)	1.2693 (0.0589)	0.2878 (0.0012)	0.6817 (0.0271)	0.9683 (0.0357)	0.1988 (0.0026)
$\xi_i$	0.1034 (0.0023)	0.0297 (0.0000)	-0.0340 (0.0072)	0.0067 (0.0001)	0.0015 (0.0000)	0.0046 (0.0000)
$\eta_i$	0.0872 (0.0018)	0.1892 (0.0016)	0.3769 (0.0084)	0.0017 (0.0001)	0.0049 (0.0002)	0.2704 (0.0053)
$\kappa_i$	0.0859 (0.0005)	0.0193 (0.0004)	-0.0095 (0.0003)	0.1164 (0.0025)	0.0053 (0.0001)	0.1222 (0.0028)
$\gamma_i$	0.0181 (0.0002)	0.0488 (0.0003)	0.02079 (0.0119)	0.1956 (0.0020)	0.0449 (0.0002)	0.2114 (0.0028)
$\rho_i^{x\sigma}$	-0.4126 (0.0020)	-0.4416 (0.0006)	0.0432 (0.0033)	0.0074 (0.0002)	-0.4376 (0.0062)	0.0127 (0.0002)
$\varpi_i$	0.1145 (0.0023)	0.4785 (0.0153)	-0.0338 (0.0020)	0.6270 (0.0858)	-0.0130 (0.0004)	0.2238 (0.0107)
$\varpi_i^\sigma$	-0.0090 (0.0001)	-0.0917 (0.0015)	0.7518 (0.0119)	0.4740 (0.0057)	-2.1178 (0.0067)	0.7654 (0.0325)
$f_0$	7.305		26.723		3.675	
$\sigma_f$	2.00%		0.95%		2.00%	
$\sigma_o$	6.87%		1.96%		6.95%	
log L	-60139		-78143		-62160	
AIC	120316		156324		124358	
RMSE Futures	2.8635%		1.8292%		1.6565%	
RMSE Imp.Vol. 4mth	1.2075%		6.9984%		1.7917%	
RMSE Imp.Vol. 12mth	1.4787%		4.4972%		1.7302%	
RMSE Imp.Vol. 2yr	1.9997%		4.1097%		1.4997%	
RMSE Imp.Vol. 3yr	2.3210%		3.3924%		0.9885%	
RMSE Imp.Vol. 4yr	2.4634%		3.9361%		1.0235%	
RMSE Imp.Vol. 5yr	2.6908%		3.8580%		1.3064%	

Table 3: **Parameter estimates of the two-dimensional model for crude oil futures and options, when  $\rho_i^{xr} = \rho_i^{r\sigma} = 0$ .** The table displays the maximum-likelihood estimates (with the standard errors in parenthesis) of the two-dimensional model specifications for the three two-year periods, namely, August 2005 – July 2007, July 2007–June 2009 and January 2011 – December 2012. Here  $f_0$  is the time-homogenous futures price at time 0, namely  $F(0, t) = f_0, \forall t$ . The quantities  $\sigma_f$  and  $\sigma_o$  are the standard deviations of the log futures prices measurements errors and the option price measurement errors, respectively. We normalise the long run mean of the volatility process,  $\bar{\sigma}_i$ , to one to achieve identification.

	Period 1			Period 2			Period 3		
	Aug 2005–Jul 2007			Jul 2007–Jun 2009			Jan 2011–Dec 2012		
	i=1	i=2	i=3	i=1	i=2	i=3	i=1	i=2	i=3
$\xi_{0i}$	0.0295 (0.0015)	0.3164 (0.0005)	2.0226 (0.0405)	0.2759 (0.0035)	0.5926 (0.0035)	1.0962 (0.0049)	0.6734 (0.0062)	0.3183 (0.0082)	0.0143 (0.0004)
$\xi_i$	0.2710 (0.0090)	0.0334 (0.0004)	0.0068 (0.0001)	0.4128 (0.0069)	0.0047 (0.0002)	0.01078 (0.0008)	0.0037 (0.0000)	0.0104 (0.0004)	0.0519 (0.0016)
$\eta_i$	0.1748 (0.0004)	0.3526 (0.0200)	0.0209 (0.0003)	0.2361 (0.0025)	0.4044 (0.0035)	0.0152 (0.0011)	0.0053 (0.0000)	0.2972 (0.0006)	0.0597 (0.0002)
$\kappa_i$	0.0287 (0.0005)	0.0670 (0.0001)	-0.0319 (0.0010)	0.0068 (0.0002)	0.0263 (0.0002)	-0.0150 (0.0002)	0.0105 (0.0002)	0.1606 (0.0052)	0.0025 (0.0000)
$\gamma_i$	-0.0311 (0.0009)	1.0054 (0.0254)	0.0064 (0.0002)	-0.0487 (0.0009)	0.6015 (0.0180)	0.00244 (0.0003)	-0.0516 (0.0005)	-0.2799 (0.0005)	0.0750 (0.0038)
$\rho_i^{\sigma\sigma}$	-0.0894 (0.0024)	-0.0962 (0.0040)	-0.1767 (0.0064)	0.0946 (0.0031)	-0.2077 (0.0009)	0.1643 (0.0113)	0.4008 (0.0005)	0.0015 (0.0000)	0.0159 (0.0003)
$\varpi_i$	-0.3211 (0.0106)	1.6172 (0.0268)	-16.4285 (0.3864)	0.9070 (0.0076)	10.3959 (0.0268)	-9.0959 (0.0172)	0.2504 (0.0025)	0.0037 (0.0000)	-0.0586 (0.0028)
$\varpi_i^\sigma$	1.1087 (0.0030)	0.1652 (0.0001)	-1.8788 (0.0349)	2.3040 (0.0541)	-0.2619 (0.0035)	0.0675 (0.0028)	-0.2761 (0.0043)	0.8139 (0.0091)	-1.0503 (0.0177)
$f_0$		9.0851			23.5172			5.5110	
$\sigma_F$		1.27%			0.45%			1.66%	
$\sigma_O$		3.15%			0.65%			11.34%	
log L		-38670			-65291			-31096	
AIC		77394			130636			62246	
RMSE Futures		1.2995%			1.1421%			1.2535%	
RMSE Imp.Vol. 4mth		1.1827%			3.6883%			1.7927%	
RMSE Imp.Vol. 12mth		1.6173%			2.8619%			1.7064%	
RMSE Imp.Vol. 2yr		1.9285%			3.0689%			1.5372%	
RMSE Imp.Vol. 3yr		2.1934%			2.9486%			1.0613%	
RMSE Imp.Vol. 4yr		2.5390%			2.8921%			0.8927%	
RMSE Imp.Vol. 5yr		2.9145%			2.7094%			0.9096%	

Table 4: **Parameter estimates of the three-dimensional model for crude oil futures and options, when  $\rho_i^{\sigma\sigma} = \rho_i^{\sigma\sigma} = 0$ .** The table displays the maximum-likelihood estimates (with the standard errors in parenthesis) of the three-dimensional model specifications for the three two-year periods, namely, August 2005 – July 2007, July 2007–June 2009 and January 2011 – December 2012. Here  $f_0$  is the time-homogenous futures price at time 0, namely  $F(0, t) = f_0, \forall t$ . The quantities  $\sigma_f$  and  $\sigma_o$  are the standard deviations of the log futures prices measurements errors and the option price measurement errors, respectively. We normalise the long run mean of the volatility process,  $\bar{\sigma}_i$ , to one to achieve identification.

		Period Aug 2005–Jul 2007						
		$\rho^{r\sigma}$						
		-0.50	-0.30	-0.15	0	0.15	0.30	0.50
$\rho^{xr}$	-0.50	10.967	10.922	10.889	10.857	10.825	10.794	10.753
	-0.30	10.823	10.78	10.748	10.717	10.686	10.656	10.617
	-0.15	10.717	10.675	10.644	10.613	10.583	10.554	10.516
	0	10.611	10.57	10.54	10.51	10.481	10.453	10.416
	0.15	10.506	10.466	10.437	10.408	10.38	10.353	10.317
	0.30	10.403	10.364	10.335	10.307	10.28	10.253	10.218
	0.50	10.266	10.228	10.201	10.174	10.148	10.122	10.089

		Period Jan 2011–Dec 2012						
		$\rho^{r\sigma}$						
		-0.50	-0.30	-0.15	0	0.15	0.30	0.50
$\rho^{xr}$	-0.50	28.605	28.6	28.595	28.591	28.587	28.583	28.577
	-0.30	28.449	28.443	28.439	28.435	28.43	28.426	28.421
	-0.15	28.332	28.326	28.322	28.318	28.314	28.309	28.304
	0	28.215	28.209	28.205	28.201	28.197	28.193	28.187
	0.15	28.098	28.093	28.089	28.085	28.081	28.077	28.071
	0.30	27.982	27.977	27.973	27.969	27.965	27.961	27.955
	0.50	27.828	27.823	27.819	27.815	27.811	27.807	27.802

Table 5: **Call Option prices for varying correlation coefficients.** Futures=100, strike=100, maturity=4000 days. We denote for example  $\rho^{xr} = -0.50$  to mean  $\rho_1^{xr} = \rho_2^{xr} = \rho_3^{xr} = -0.50$  and similarly for  $\rho^{r\sigma}$ .

	Period 1		Period 2		Period 3	
	Aug 2005–Jul 2007		Jul 2007–Jun2009		Jan 2011–Dec 2012	
	$i = 1$	$i = 2$	$i = 1$	$i = 2$	$i = 1$	$i = 2$
$\rho_i^{xr}$	-0.6166	-0.3014	-0.0102	-0.6640	0.5420	-0.6409
	(0.0133)	(0.0045)	(0.0007)	(0.0597)	(0.0061)	(0.0099)
log L	-59977		-73689		-61299	
AIC	119996		147420		122640	
RMSE Futures	2.8408%		1.7729%		1.6563%	
RMSE Imp.Vol. 4mth	1.2174%		6.8423%		1.7916%	
RMSE Imp.Vol. 12mth	1.4736%		4.2386%		1.7233%	
RMSE Imp.Vol. 2yr	1.9605%		4.1089%		1.4987%	
RMSE Imp.Vol. 3yr	2.2144%		3.3136%		0.9987%	
RMSE Imp.Vol. 4yr	2.2947%		3.6888%		1.0082%	
RMSE Imp.Vol. 5yr	2.4606%		3.7230%		1.2727%	

Table 6: **Parameter estimates of  $\rho_i^{xr}$  for the two-dimensional model for crude oil futures and options.** The table displays the maximum-likelihood estimates of  $\rho_i^{xr}$  (with the standard errors in parenthesis) of the two-dimensional model specifications for the three two-year periods, namely, August 2005 – July 2007, July 2007–June 2009 and January 2011 – December 2012.

	Period 1			Period 2			Period 3		
	Aug 2005–Jul 2007			Jul 2007–Jun2009			Jan 2011–Dec 2012		
	i=1	i=2	i=3	i=1	i=2	i=3	i=1	i=2	i=3
$\rho_i^{xr}$	-0.6222 (0.0013)	-0.3822 (0.0110)	-0.5444 (0.0256)	-0.0245 (0.0003)	-0.1426 (0.0062)	-0.0951 (0.0009)	-0.3167 (0.0052)	0.59167 (0.0005)	-0.4583 (0.0002)
log L		-35893			-63557			-31372	
AIC		71846			127174			62804	
RMSE Futures		1.3067%			1.1125%			1.2452%	
RMSE Imp.Vol. 4mth		1.1792%			3.6571%			1.7925%	
RMSE Imp.Vol. 12mth		1.6091%			2.9659%			1.7078%	
RMSE Imp.Vol. 2yr		1.8815%			3.0873%			1.5345%	
RMSE Imp.Vol. 3yr		2.0380%			2.5734%			1.0497%	
RMSE Imp.Vol. 4yr		2.3487%			2.3525%			0.8897%	
RMSE Imp.Vol. 5yr		2.6539%			2.1850%			0.9139%	

Table 7: **Parameter estimates of  $\rho_i^{xr}$  for the three-dimensional model for crude oil futures and options.** The table displays the maximum-likelihood estimates of  $\rho_i^{xr}$  (with the standard errors in parenthesis) of the three-dimensional model specifications for the three two-year periods, namely, August 2005 – July 2007, July 2007–June 2009 and January 2011 – December 2012.

	Period 1: Aug 05–Jul 07				Period 2: Jul 07–Jun 09				Period 3: Jan 11–Dec 12			
	IS		OS		IS		OS		IS		OS	
	Sto	Det	Sto	Det	Sto	Det	Sto	Det	Sto	Det	Sto	Det
4mth	1.18%	1.18%	1.30%	1.34%	3.24%	3.27%	4.98%	5.21%	1.75%	1.75%	1.88%	1.88%
12mth	1.61%	1.63%	2.43%	2.59%	3.05%	3.22%	2.59%	2.70%	1.72%	1.72%	1.64%	1.67%
2year	1.90%	1.98%	2.38%	2.58%	3.24%	3.49%	2.42%	2.50%	1.53%	1.51%	1.59%	1.65%
3year	2.39%	2.68%	2.50%	3.15%	2.87%	3.51%	1.21%	1.25%	1.27%	1.18%	1.43%	1.60%
4year	2.75%	3.13%	2.87%	3.77%	2.65%	3.52%	0.86%	1.03%	1.02%	0.89%	1.54%	1.81%
5year	3.11%	3.58%	3.23%	4.06%	2.42%	3.42%	1.18%	1.42%	1.04%	0.89%	1.91%	2.26%

Table 8: **RMSE of implied volatility.** The table displays the RMSE between the observed implied volatility and the implied volatility from the estimated model for different maturities. In-sample (IS) and out-of-sample (OS) analysis is included.

	Period 1: Aug 05–Jul 07				Period 2: Jul 07–Jun 09				Period 3: Jan 11–Dec 12			
	IS		OS		IS		OS		IS		OS	
	Sto	Det	Sto	Det	Sto	Det	Sto	Det	Sto	Det	Sto	Det
1year	4.24%	4.35%	4.98%	5.21%	8.72%	9.53%	4.32%	4.56%	5.49%	5.55%	4.81%	4.80%
2year	4.65%	5.04%	4.98%	5.20%	9.18%	9.82%	4.15%	4.40%	6.55%	6.45%	4.67%	4.81%
3year	4.98%	4.99%	4.19%	4.75%	7.24%	7.50%	2.21%	2.37%	4.78%	4.58%	3.64%	3.92%
4year	5.63%	5.96%	4.11%	5.07%	8.16%	9.06%	1.70%	2.14%	3.49%	2.95%	3.55%	3.89%
5year	6.81%	6.79%	4.93%	6.48%	8.85%	9.75%	2.29%	2.95%	4.07%	3.14%	4.67%	5.08%

Table 9: **Maximum absolute difference in implied volatilities.** The table displays the maximum absolute difference between the volatility implied by the market crude oil option prices and the model implied volatility (from the deterministic interest rates model and the stochastic interest rates model) observed in the in-sample (IS) period and the out-of-sample (OS) period, over a range of maturities.

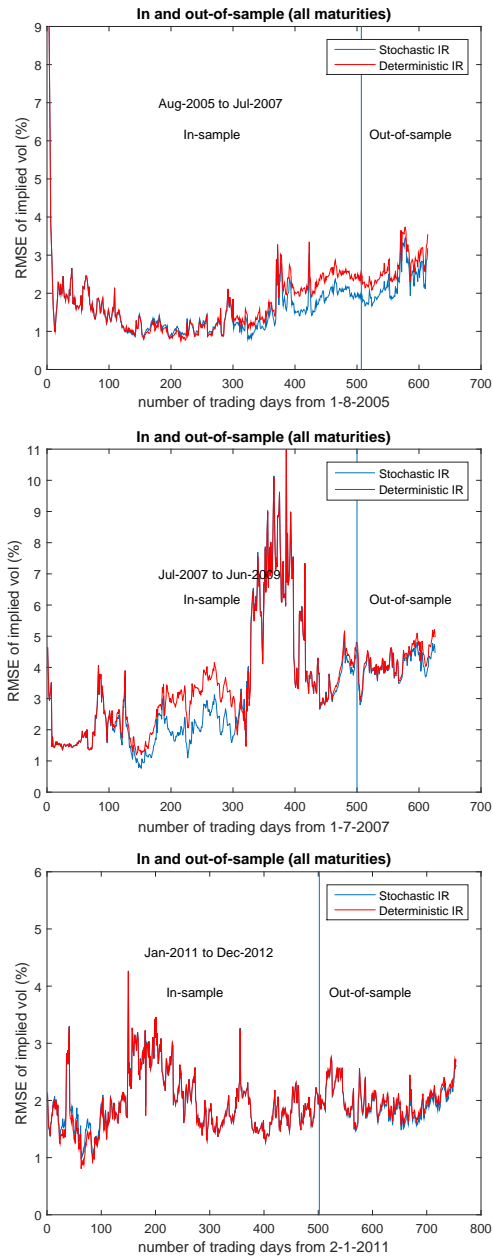


Figure 6: **RMSE of implied volatility for all maturities.** The graph displays the average of the daily RMSE of the implied volatility across all maturities. The top panel displays the fitting between August 2005 – July 2007 (and August 2007–December 2007 for the out-of-sample), the middle panel displays the fitting between July 2007 – June 2009 (and July 2009–December 2009 for the out-of-sample) and the bottom panel displays the fitting between January 2011 – December 2012 (and January 2013–June 2013 for the out-of-sample). The RMSE is defined to be between the volatility implied by observed market prices of the crude oil options and the implied volatility from the estimated model.

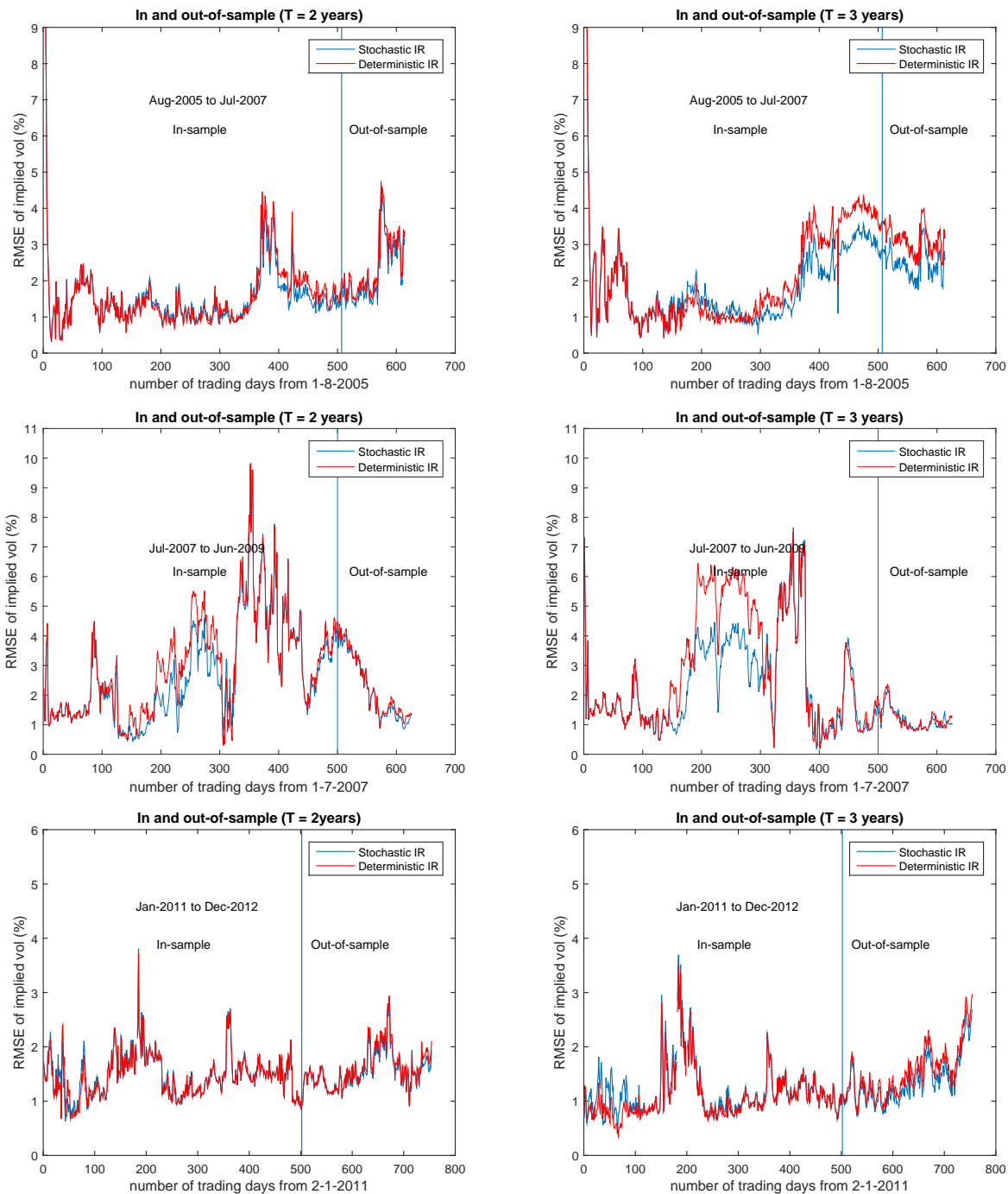


Figure 7: **RMSE of implied volatility with 2-year and 3-year maturities.** The graphs show the daily RMSE of the implied volatility (the RMSE between the observed implied volatility and the implied volatility from the estimated model). The top panel displays the fitting between August 2005 – July 2007 (and August 2007–December 2007 for the out-of-sample), the middle panel displays the fitting between July 2007 – June 2009 (and July 2009–December 2009 for the out-of-sample) and the bottom panel displays the fitting between January 2011 – December 2012 (and January 2013–June 2013 for the out-of-sample). The graphs on the left show the RMSE of options fitted to December contracts which are in the second year, while the graphs on the right show the RMSE of options fitted to December contracts which are in the third year from the beginning of the sample period.



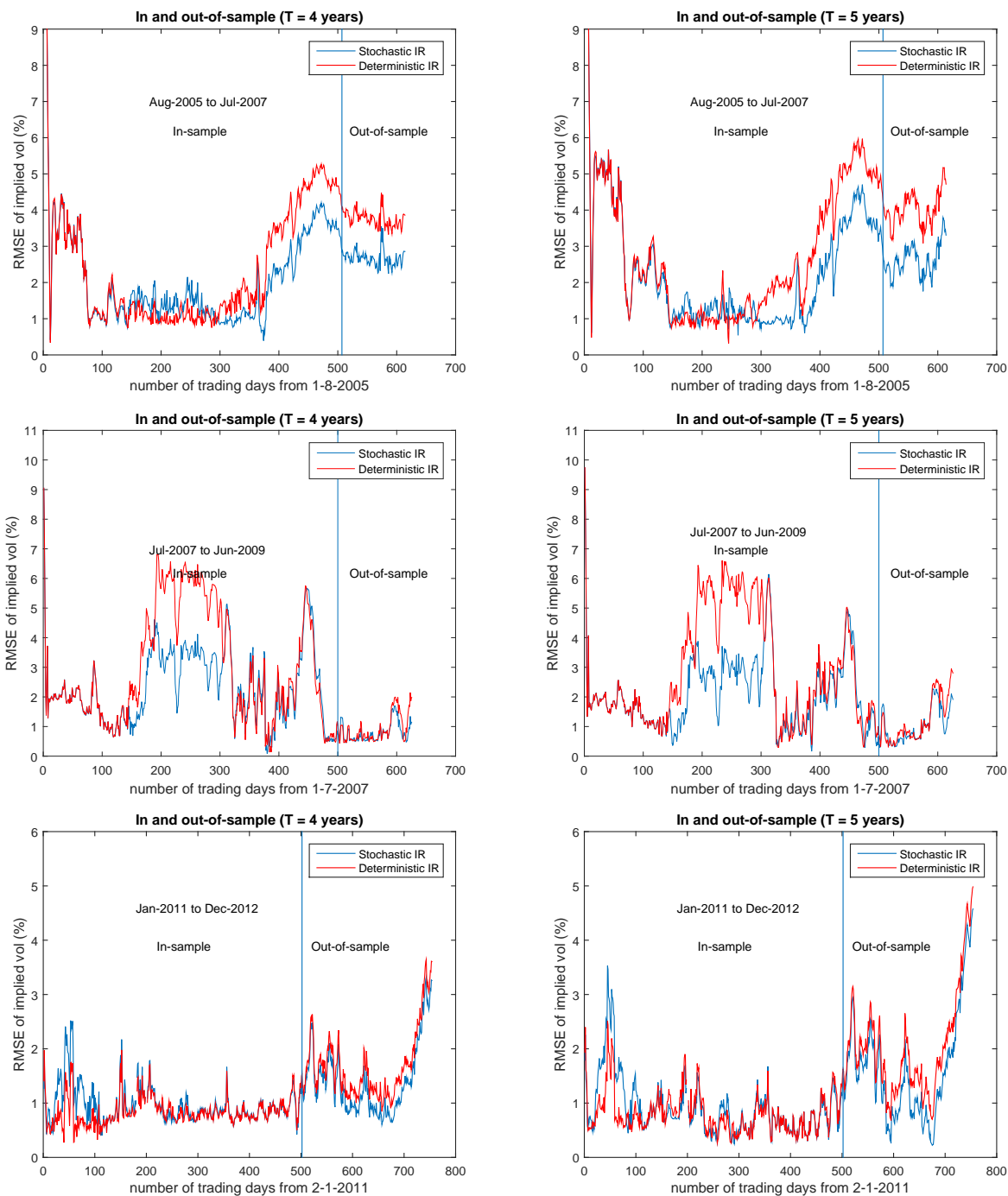


Figure 8: **RMSE of implied volatility with 4-year and 5-year maturities.** The graphs show the daily RMSE of the implied volatility (the RMSE between the observed implied volatility and the implied volatility from the estimated model). The top panel displays the fitting between August 2005 – July 2007 (and August 2007–December 2007 for the out-of-sample), the middle panel displays the fitting between July 2007 – June 2009 (and July 2009–December 2009 for the out-of-sample) and the bottom panel displays the fitting between January 2011 – December 2012 (and January 2013–June 2013 for the out-of-sample). The graphs on the left show the RMSE of options fitted to December contracts which are in the fourth year, while the graphs on the right show the RMSE of options fitted to December contracts which are in the fifth year from the beginning of the sample period.