# Maximum Likelihood Estimation of Stochastic Volatility and Pricing Derivatives in Commodity Markets

Damien Lee

School of Banking and Finance
The University of New South Wales

A thesis submitted to the University of New South Wales in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

2010

**COPYRIGHT STATEMENT**



Signed …….....

Date …….....

**AUTHENTICITY STATEMENT**

'I certify that the Library deposit digital copy is a direct equivalent of the final officially approved version of my thesis. No emendation of content has occurred and if there are any minor variations in formatting, they are the result of the conversion to digital format.'

Signed …….....

Date …….....

**ORIGINALITY STATEMENT**

'I hereby declare that this submission is my own work and to the best of my knowledge it contains no materials previously published or written by another person, or substantial proportions of material which have been accepted for the award of any other degree or diploma at UNSW or any other educational institution, except where due acknowledgement is made in the thesis. Any contribution made to the research by others, with whom I have worked at UNSW or elsewhere, is explicitly acknowledged in the thesis. I also declare that the intellectual content of this thesis is the product of my own work, except to the extent that assistance from others in the project's design and conception or in style, presentation and linguistic expression is acknowledged.'

Signed ……..

Date ……..

**Abstract**

Financial markets worldwide have grown rapidly over the last few decades and so have the number of modelling approaches to analyse and price financial assets. With the emergence of more complex models, emphasis has also been placed on calibrating the models to market data. This thesis consists of three distinct components which investigate the estimation of financial models and the pricing of derivatives. The focus of these studies is stochastic volatility models and the crude oil futures market.

The first component investigates discrete-time stochastic volatility models by comparing the estimation performance of three maximum-likelihood procedures. The analysis is conducted empirically on the fixed income and crude oil markets and also tests the validity of different stochastic volatility model specifications in-sample and out-of-sample. The study finds that the choice of estimation procedure is important if conditional volatility estimates are required. Also, a traditional AR(1) specification for the log-variance is sufficient in the fixed income and crude oil markets.

The second component introduces a three-factor short/long factor commodity model which allows for mean-reversion in spot prices, expected increases in long-term prices and a time-varying market price of risk. The model is able to accurately capture the term structure of futures prices in the crude oil futures market with evidence suggesting that risk premiums are time-varying. Using the cross-section of futures prices we estimate a time-series of the market price of risk implied by the model. We find that the risk premiums in the crude oil market are driven by the same risk factors as equity and bond markets.

In the final component, the short/long factor model is extended to incorporate both jumps and stochastic volatility. Semi-analytical solutions of futures and European option prices are derived for the model. The futures and option pricing performance is compared with nested specifications. The

empirical results demonstrate that although introducing jumps or stochastic volatility does not impact futures pricing much, they are required for option pricing applications. When fitting the implied volatility surface of crude oil futures options, stochastic volatility is required to fit implied volatility over the maturity and moneyness dimensions but jumps are required when fitting the steep volatility smiles exhibited by short-term options.

# Acknowledgements

As with all large undertakings, I have a number of people to thank, from those who got me started on this journey to those who saw me through to the end. Firstly, I'd like to thank my supervisor Associate Professor Ramaprasad Bhar. You have always provided invaluable insight during our discussions and provided me with much needed direction in my research. I appreciate the support you gave me during my candidature and I am especially thankful for giving me the opportunity to work under you despite little research experience.

I want to thank Udara Peiris for helping me to see the limitations of my previous career path and that there are better opportunities available. Without your introduction to some of these opportunities, I probably would not have even started my PhD.

I would also like to thank Dr David Colwell who was always kind enough to give me time for a chat. Your help in checking my work was extremely appreciated even though you had no reason to do it. To Dr Li Yang and Associate Professor Toan Pham, thank you for providing me with the opportunity to get some teaching experience. The experience was not only valuable but also assisted financially.

I am also grateful to Associate Professor Pradeep Ray who provided me with financial assistance for part of my candidature. Although I did not pursue a path related to ontologies, I found the discussions in the area to be very interesting.

I also want to thank my family who have been very patient and providing me with plenty of support. Finally I would like to thank Fiona for giving me so much support,

encouragement and love. Through the highs and the lows, you were always there for me. Without you, this would not have been possible.

# Contents

# List of Tables

# List of Figures

# Chapter 1

# Introduction

Financial modelling has been at the forefront of the finance literature as both academics and practitioners alike have attempted to understand and predict the behaviour of financial markets. The rapid growth of financial markets worldwide over the last few decades has fueled the growth in the literature. From the introduction of the model of stock prices by Black and Scholes (1973) and Merton (1973) (BS-M), the financial modelling literature has grown tremendously. As the assumptions behind the BS-M modelling approach have proven too simple to accurately model asset prices in practice, the literature has introduced dynamics to models which reflect the empirical bahaviour of asset prices. This has resulted in numerous modelling approaches which attempt to explain many of the "stylized facts" or characteristics exhibited by asset prices. However, in order to capture many of these characteristics, the models introduced have increased in complexity. Consequently, this has come at the cost of an increased difficulty when calibrating the models to market data and pricing derivative securities based on the models.

However, the literature has also expanded in these areas with advances in estimation techniques and techniques for pricing derivatives. Progress in these areas have focussed on equity, fixed-income, and foreign exchange markets, but the literature has tended to lag behind for commodities due to the relatively smaller size of commodity markets. Also, as commodities are consumption assets, they tend to

exhibit different characteristics to assets in financial markets. This has spawned modelling approaches slightly removed from the wider finance literature. However, commodities still show behaviour similar to financial markets which allows the same techniques introduced in other markets to be applied in commodity markets. This thesis contributes to the literature in this regard by considering some of the advances in modelling and estimation in financial markets and applying them to commodity markets. The focus of this thesis is on stochastic volatility models and the empirical applications mainly consider crude oil as it is the world's most liquid commodity. However, other areas in the finance literature are also covered.

The first component of the thesis in Chapter 2 investigates discrete-time stochastic volatility (SV) models. The discrete-time SV modelling literature has spawned from the model introduced by Taylor (1986) where the log-variance is modelled as an autoregressive process. They have proven to be a popular alternative to GARCH for modelling time-varying volatility and have been found to be able to capture the behaviour of volatility more accurately. The study concentrates on two areas that have received little attention in the SV literature, the fixed income market and crude oil markets. The study considers the suitability of SV models for these markets and looks at what difference the estimation procedure has on the choice of model. To achieve this, three maximum likelihood estimation procedures are employed to estimate alternative SV model specifications. The three procedures include quasi-maximum likelihood (QML), Monte Carlo likelihood (MCL) and particle filtering (PF) where the latter two approaches apply importance sampling techniques. Using a number of diagnostic tests, the model specifications and conditional volatility estimates are evaluated both in-sample and out-of-sample. The traditional tests of predictive accuracy such as root mean-squared errors or mean absolute errors for SV models are problematic given that volatility is unobserved and a suitable proxy for volatility is difficult to obtain. The study considers an alternative diagnostic test

to compare the performance of SV models based on the distribution of the asset returns.

The second component in Chapter 3 considers a three-factor short/long factor model for the modelling of commodity prices. The model consists of two short-term factors which represent the short term deviation in spot prices, and a long-term factor which represents an equilibrium price level. The model captures a number of features exhibited by commodities such as mean-reversion in short-term prices, expected long-term increase in prices, the Samuelson effect and also allows for time-varying risk premiums. Using a panel data set of crude oil futures data, the performance of the three-factor model and a two-factor nested specification are analysed. It is found that the three-factor model is able to accurately capture the term structure of futures prices with evidence suggesting that risk premiums in the crude oil market are time-varying. We then extract a time-series of the model implied risk premiums. This allows for an analysis of risk premiums where it is determined whether they can be explained by the same risk factors as equity and bond markets.

Chapter 4 comprises the third component where an extension to the short/long factor model is introduced which incorporates jumps in prices and stochastic volatility. Similar to financial markets, commodity prices have been empirically observed to exhibit large movements in prices resulting in fat tails of return distributions, time-varying volatility and volatility clustering. Although the Gaussian short/long model is able to accurately recover the term structure of futures prices, it cannot explain the distribution of returns and option pricing performance is likely to be poor due to the assumption of constant volatility. The study introduces a stochastic volatility jump-diffusion model to capture the behaviour exhibited by commodities and derives semi-closed form[1] solutions to futures and option prices. Using crude oil

[1]Semi-closed form as they are solvable up to the solution of a system of ordinary differential equations.

futures and options data, the study investigates the impact of introducing either jumps or stochastic volatility to the model dynamics. Although the extra dynamics have little effect on the futures term structure, crude oil prices clearly exhibit non-Gaussian behaviour. Also, it is well known that options in equity markets exhibit implied volatility "smiles" and "skew" and the same behaviour has been found in some recent studies of crude oil futures options. This study investigates the option pricing performance and determines whether jumps or stochastic volatility are necessary for modelling in the crude oil futures market.

Finally, Chapter 5 summarises the thesis and provides future avenues of research.

## 1.1 State Space Models

In each component of the thesis, estimation is conducted under a state space modelling framework. State space models have become popular in financial modelling as it allows filtering and estimation using tools which already exist in engineering disciplines. They typically consist of a measurement or observation equation and a state or transition equation. The state equation describes the evolution of a latent state variable from past states to future states. The measurement equation describes the relationship of the latent state variable to the observations. Throughout this thesis, we will denote the vector of observations at each time by $y_t$, for $t = 1, \ldots, T$, and the entire set of observations by $y = (y_1, \cdots, y_T)'$. Similarly, the vector of states are represented by $x_t$ and the entire set of states by $x = (x_0, \cdots, x_T)'$. The simplest case of a state space model is a linear Gaussian state space model and is described

by the following equation,

$$\begin{aligned} y_t &= Z_t x_t + d_t + \epsilon_t, \\ x_t &= T_t x_{t-1} + c_t + \eta_t, \end{aligned} \tag{1.1}$$

where for $t = 1, \ldots, T$, $y_t \in \mathbb{R}^n$, $x_t \in \mathbb{R}^m$ and the error terms, $\epsilon_t \in \mathbb{R}^n$ and $\eta_t \in \mathbb{R}^m$ are Gaussian with $\epsilon_t \sim N(0, H_t)$ and $\eta_t \sim N(0, Q_t)$. The other variables, $Z_t \in \mathbb{R}^n \times \mathbb{R}^m$, $d_t \in \mathbb{R}^n$, $T_t \in \mathbb{R}^m \times \mathbb{R}^m$ and $c_t \in \mathbb{R}^m$ describe the dynamics of the state space system. For the linear Gaussian model, the well-known Kalman filter and smoother can be applied for model inference such as estimating the posterior mean or estimating the likelihood function. The Kalman filter is described in the next section as it is used throughout this thesis.

In many situations considered in finance, the models are not linear Gaussian models. A more general state space model which does not assume linearity of either the observation or state equations can be expressed as follows

$$\begin{aligned} y_t &= h(x_t, \epsilon_t), \\ x_t &= f(x_{t-1}, \eta_t), \end{aligned} \tag{1.2}$$

where the functions are the mappings $h : \mathbb{R}^m \times \mathbb{R}^n \mapsto \mathbb{R}^n$ and $f : \mathbb{R}^m \times \mathbb{R}^m \mapsto \mathbb{R}^m$. In this thesis, it is assumed that the state is a Markov process and that the observations are independent given the states. This means that the distributions of $\epsilon_t$ and $\eta_t$ also satisfy those requirements. Under this specification the observations and states can be both nonlinear and non-Gaussian. However, when the model is nonlinear and/or non-Gaussian, estimation is generally not analytically tractable and various methods have been introduced in the literature to tackle the issue. Chapter 2 will investigate a few of these procedures in the case of maximum likelihood estimation for SV models.

## 1.2 The Kalman Filter

Originally developed by Kalman (1960), the Kalman filter aims to estimate the true values of an underlying model given noisy measurements or observations. It is the most optimal filter for linear Gaussian state space models in the sense that it produces the minimum mean-square estimator (MMSE). The procedure is applied extensively in this thesis and is introduced here for future reference. The algorithm for performing the Kalman filter has a number of references in the literature, however for a thorough treatment of its application to financial modelling, readers are directed to Harvey (1989). Here we only describe the linear Kalman filter recursions. Given the initial state $x_0 \sim N(\bar{x}_0, P_0)$ and iterating over $t = 1, \ldots, T$, the Kalman filter consists of a prediction step,

$$\bar{x}_{t|t-1} = T_t \bar{x}_{t-1|t-1} + c_t, \tag{1.3}$$

$$P_{t|t-1} = T_t P_{t-1|t-1} T_t^{'} + Q_t, \tag{1.4}$$

and an update step,

$$\bar{y}_{t|t-1} = Z_t \bar{x}_{t|t-1} + d_t, \tag{1.5}$$

$$v_t = y_t - \bar{y}_{t|t-1}, \tag{1.6}$$

$$F_t = Z_t P_{t|t-1} Z_t^{'} + H_t, \tag{1.7}$$

$$K_t = P_{t|t-1} Z_t^{'} F_t^{-1}, \tag{1.8}$$

$$\bar{x}_{t|t} = \bar{x}_{t|t-1} + K_t v_t, \tag{1.9}$$

$$P_{t|t} = (I - K_t Z_t) P_{t|t-1}. \tag{1.10}$$

Generally, parameter estimation is conducted via maximum likelihood estimation. For a linear Gaussian state space model, an analytical solution for the likelihood

function is known. Given the parameter set of the model, $\Theta$, the marginal log-likelihoods at each time-step are given by

$$\log l_t(\Theta) = \log p(y_t|y_{1:t-1};\Theta) = -\frac{1}{2}\left(n\log(2\pi) + \log|F_t| + v_t F_t^{-1} v_t^{'}\right). \tag{1.11}$$

where the use of $\Theta$ in the equation emphasizes the dependence of the marginal log-likelihoods on the parameters of the model[2]. The log-likelihood is therefore given by

$$\log L(\Theta) = \sum_{t=1}^{T} \log l_t(\Theta). \tag{1.12}$$

The parameter estimates are obtained by maximizing the likelihood function.

[2]The parameter set depends on the particular model being studied. As no model has been specified as yet, $\Theta$ is left undefined.

# Chapter 2

# Comparing Maximum Likelihood Estimation of Stochastic Volatility: The case of short-term interest rates and crude oil

## 2.1 Introduction

In most of the modelling approaches in finance, it is generally accepted that non-Gaussian dynamics are needed in asset pricing models. Assuming more complex distributions allows researchers to develop theoretical models that are able to replicate many of the empirical properties observed in financial markets. For instance, the distribution of financial returns have been observed to have fatter tails than a Gaussian distribution as well as exhibiting time-varying variance. However, a critical issue then becomes how to calibrate the model to real data and in the case of latent variable models, how to estimate the underlying latent state variables. Generally this issue is addressed by the choice of estimation procedure although for more complex models this is not always straightforward. The choice of estimation procedure can be complicated by not only how accurate the procedure is but also the ease of implementation and its computational efficiency.

One area of financial modelling that has spawned many different estimation procedures with varying degrees of efficiency and speed is stochastic volatility modelling. Stochastic volatility (SV) models in discrete time have been traditionally used to model time-varying volatility as an alternative to generalized autoregressive conditional heteroscedasticity (GARCH) models introduced by Bollerslev (1986). In SV models, volatility is defined as a random, unobserved latent variable of an underlying return series, where the log of the squared residuals is most commonly modelled as a first-order autoregression. SV models are attractive to researchers as they have been found to capture the empirical properties observed in financial time-series better than GARCH models. However, estimation is not straightforward, which has resulted in the introduction of a number of procedures in the literature to estimate SV models.

Estimating stochastic volatility models has been the subject of a number of surveys in the literature. A fairly comprehensive review of the recent literature can be found in Part II of Anderson et al. (2009). Broto and Ruiz (2004) also provides a survey of some of the techniques available for estimation in SV models and we discuss some of the main techniques here. The earliest techniques include the generalized method of moments (GMM) approach of Melino and Turnbull (1990) or the quasi-maximum likelihood (QML) approach of Harvey et al. (1994). These approaches provide relatively less efficient estimates than most approaches proposed in the literature but have been popular due to their computational efficiency and simplicity. More accurate techniques can be split between Monte Carlo Markov Chain (MCMC) methods such as the Jacquier et al. (1994) method or simulated maximum likelihood procedures that have followed from the Danielsson (1994) approach which employs importance sampling. Two approaches that use importance sampling techniques include the Monte Carlo Likelihood (MCL) of Sandmann and Koopman (1998) and particle filters (PF). Whilst MCMC procedures are quite popular in the literature

for SV modelling, they are known to be very computationally demanding. On the other hand, Broto and Ruiz (2004) found that in terms of estimation accuracy, some simulated maximum likelihood procedures performed comparably with MCMC techniques and tend to be much less computationally intensive. In this chapter, we compare three maximum likelihood estimation procedures in the SV modelling literature, QML, MCL and PF.

The motivation for the choice of estimation procedures used in the comparison is due to a number of reasons. With QML, despite the known inefficiency of its estimates, has been quite popular due to its ease of implementation and fast computational time. For the MCL procedure, whilst less popular than MCMC, has been found by Sandmann and Koopman (1998) and Broto and Ruiz (2004) to exhibit similar performance to MCMC in finite samples whilst being much more computationally efficient. Particle filters have become increasingly popular in financial applications for estimating nonlinear and/or non-Gaussian models and have also been found to have comparable performance to MCMC methods in financial modelling. Furthermore, implementation for each of these procedures is not overly difficult and all three are easily modified for applications other than SV modelling. Broto and Ruiz (2004), Koopman and Lee (2004) and Sapp (2009) also provide a comparison of estimation procedures for SV models, however no study has compared particle filters.

To compare the estimation procedures, we firstly conduct a brief finite-sample study of the performance of the estimators. We then consider an empirical application in two areas that have received relatively little attention within the stochastic volatility literature, fixed income and commodity markets. It has been well established that most asset markets exhibit time-varying volatility and both fixed-income and commodity markets are no exception. In terms of the literature on discrete-time SV

models in these two areas, Ball and Torous (1999) used the framework of Chan et al. (1992) (hereafter CKLS) and showed that the volatility of short-term interest rates depends not only on the level of interest rates but also exhibits stochastic volatility. This was further confirmed by Sapp (2009) who estimated a similar model to Ball and Torous (1999) by comparing the performance of a GMM estimator and the QML approach. Other papers (e.g. Smith (2002), Kalimipali and Susmel (2004), Sun (2005), Vo (2009)) consider extending the SV model to incorporate Markov switching in the volatility dynamics of short-term interest rates. For the commodity market, the only paper that investigates SV models of the type considered here includes the paper by Vo (2009) who investigates stochastic volatility in the crude oil market. This chapter demonstrates estimation of stochastic volatility models for short-term interest rate and crude oil markets using the three maximum likelihood procedures, QML, MCL and PF. This study aims to determine what difference, if any, the estimation procedure has on the parameter estimates and conditional volatility estimates. It also investigates the validity of using SV models in these markets.

As mentioned previously, a standard approach to modelling stochastic volatility in the literature for equity and foreign exchange markets is an AR(1) specification of the log-variance. This study considers an extension to this model by including a moving average (MA) component in the volatility dynamics to capture any correlation with past shocks in volatility. Additionally, for short-term interest rates, it is known that macroeconomic announcements and in particular, central bank announcements, impact the volatility of interest rates as found in Das (2002). In order to capture this behaviour, the SV model for short-term interest rates is augmented to incorporate an explanatory variable for the Federal Open Market Committee (FOMC) meeting dates in the volatility dynamics.

Apart from the estimation procedures, the validity of the different specifications are also explored through various diagnostic tests which investigate both the in-sample and out-of-sample performance. As each of the procedures are filters, this facilitates estimation of the unobserved volatility process allowing us to directly compare the volatility estimates of each specification and estimation procedure. A number of empirical studies have used the absolute or squared residuals as a proxy for volatility to evaluate the performance of stochastic volatility models in empirical studies. However, the use of either of these measures as a proxy for volatility is problematic as neither can be considered accurate measures of volatility. Instead, we demonstrate alternative tests, both in-sample and out-of-sample, which examine the distributional properties of the underlying return or yield data. This study finds that the use of metrics based on the squared residuals are not reliable as the tests based on the absolute residuals gave inconsistent results especially when comparing between estimation procedures. Alternatively, for the out-of-sample tests, we adapt the Berkowitz (2001) forecast density test for SV models to determine the accuracy of the model forecasts.

This chapter is organized as follows - Section 2.2 introduces the interest rate modelling framework that incorporates stochastic volatility. Section 2.3 describes the estimation procedures used in this study. Section 2.4 presents the empirical analysis including a description of the data and the parameter estimation results. Section 2.5 conducts in-sample and out-of-sample testing for the SV model and looks at the impact of the choice of estimation procedure. Section 2.6 concludes the chapter.

## 2.2 Stochastic Volatility Model

### 2.2.1 Short-term interest rates

We consider the class of short-term interest rate models given by the following stochastic differential equation (SDE)

$$dr_t = (a + br_t)dt + \sigma_t r_t^{\gamma} dZ_t, \tag{2.1}$$

where $Z_t$ is standard Brownian motion. In the model, $a$ and $b$ determine the drift of the interest rate process and $\gamma$ is commmonly known as the elasticity of variance in the literature. Although all parameters are allowed to be freely estimated, they should satisfy some implicit assumptions that also give an additional check when estimating the model parameters. As the interest rate process is a strictly positive process, $a$ should be positive and as interest rates tend to mean-revert, $b$ should be negative. Also, due to the presence of $\gamma$, the model belongs to a class of models known as "constant elasticity of variance" (CEV) models, although here the model is augmented with stochastic volatility.

In this study, we consider the following discrete-time version of this model similar to the one introduced in CKLS which is based on an Euler-Maruyama (EM) scheme,

$$r_t - r_{t-1} = (a + br_{t-1})\Delta t + \sigma_t r_{t-1}^{\gamma} \epsilon_t, \quad \epsilon_t \sim \text{NID}(0, 1), \tag{2.2}$$

where $\Delta t$ is time between each observation. Whilst CKLS assumed $\sigma_t$ to be deterministic, a number of subsequent empirical studies found this not to be the case. As a result, a number of authors have considered extending the CKLS modelling

approach by incorporating stochastic volatility (e.g. Ball and Torous (1999), Smith (2002), Kalimipali and Susmel (2004) and Sun (2005)). In terms of estimating SV models under the CKLS framework, the usual approach involves firstly estimating (2.2) by OLS and then estimating the SV model using the residuals from the regression. In this chapter we follow a similar approach and allow the log-variance to evolve as an AR(1) process. Setting

$$y_t = \sigma_t r_{t-1}^{\gamma} \epsilon_t, \tag{2.3}$$

and letting $\sigma_t = \exp(x_t/2)$, the log-variance is given by

$$\begin{aligned} &\log \sigma_t^2 \equiv x_t = \mu(1-\phi) + \phi x_{t-1} + \eta_t, \quad \eta_t \sim \text{NID}(0, \sigma_\eta^2), \\ &x_0 \sim \text{N}(\mu, \sigma_\eta^2/(1-\phi^2)), \end{aligned} \tag{2.4}$$

where $\mu$ is the long-term mean of the log-variance, $\phi$ is the volatility persistence parameter, $\sigma_\eta$ is the volatility of volatility parameter, and $\epsilon_t$ and $\eta_t$ are assumed to be independent of each other.

### 2.2.2 Commodities

As commodities, unlike stocks or bonds, are physically traded goods, there is usually a known cost of producing them. Consequently, commodity prices generally exhibit mean-reversion toward the cost of production. Commodity prices therefore exhibit characteristics which are common to both stock prices and interest rates. By incorporating these characteristics, a model for the price of a commodity, $S_t$, is assumed to be given by the following dynamics,

$$S_t = \exp\{X_t\}, \tag{2.5a}$$

$$dX_t = (a + bX_t)dt + \sigma_t dZ_t, \tag{2.5b}$$

where $X_t = \log S_t$, $Z_t$ is standard Brownian motion and $\sigma_t$ is assumed to be stochastic. The equation for $X_t$ is similar to the process for the short-term interest rate process, $r_t$, except that the elasticity parameter $\gamma$ is set to zero so that volatility does not depend on the spot price[3]. Also, $a$ and $b$ have a similar interpretation as in the interest rate model in (2.1), except that $a$ does not need to be strictly positive, although $b$ should still be negative to reflect mean-reversion in commodity prices.

As with the interest rate process, a discrete-time version of this model is given by

$$X_t - X_{t-1} = (a + bX_{t-1})\Delta t + \sigma_t \epsilon_t,\ \epsilon_t \sim \text{NID}(0, 1). \tag{2.6}$$

In this model, $y_t$ is defined by (2.3) with $\gamma = 0$ and the log-variance is given by (2.4). Hence, the only difference between the two models is that the commodity model does not include the level of the commodity price in the volatility equation whereas the level of interest rates is included in the interest rate model.

### 2.2.3 Extensions

As an extension to the traditional AR(1) specification, this chapter also investigates specifying the log-variance as an ARMA(1,1) process. This gives a more general specification for the SV model which allows for the volatility process to be correlated not only with past levels but also past shocks in volatility. In this case, when there is

[3]Although one could use a CEV-type specification for $X_t$ with $\gamma$ unrestricted, this specification is more typically found in the literature for commodity models. However, the model was also estimated using the more general model where $\gamma$ is unrestricted. However, the $\gamma$ was found to be both close to zero and highly insignificant. Hence, only the results of the model given by (2.6) are reported where $\gamma$ is restricted to zero.

an increase (decrease) in volatility during a previous period, the increase (decrease) would be expected to persist for the current period. With the addition of the MA component, the transition equation becomes

$$\begin{aligned} x_t &= \mu(1-\phi) + \phi x_{t-1} + \theta\eta_{t-1} + \eta_t, \quad \eta_t \sim \text{NID}(0, \sigma_\eta^2), \\ x_0 &\sim \text{N}(\mu, \sigma_\eta^2(1 + (\theta+\phi)^2/(1-\phi^2))). \end{aligned} \tag{2.7}$$

Although this is a simple extension, it has not been covered in the literature for SV models. We investigate whether it can improve volatility modelling in the empirical application.

In addition, if any explanatory variables are thought to impact volatility, this can be incorporated in the SV modelling framework. For instance, some studies have considered the impact of US Federal Reserve announcements on short-term interest rates as the yields on short-term bonds tend to closely track the federal funds rate. Das (2002) considers the effect of the Federal Open Market Committee meetings on the jump probability of the federal funds rate where an increase in jump probability corresponds to an increase in the volatility of the federal funds rate. The findings suggest that the effect of 2-day FOMC meetings increased the probability of jumps occurring. Furthermore, Jones et al. (1998) and Ball and Torous (1999) found that shocks in the volatility of U.S. treasury bonds caused by macroeconomic announcements are not persistent. We attempt to model this behaviour by incorporating a dummy variable for the FOMC meetings to determine whether the activity of central banks has any impact on volatility. Denoting the FOMC dummy as $f_t$, the value is 1 if there is an FOMC meeting during the period [t,t+1] and zero otherwise. For the SV models, this can be incorporated as follows,

$$x_t = \mu(1-\phi) + \phi x_{t-1} + \kappa(1-\phi L) f_t + \eta_t, \tag{2.8}$$

$$x_t = \mu(1-\phi) + \phi x_{t-1} + \kappa(1-\phi L) f_t + \theta\eta_{t-1} + \eta_t, \tag{2.9}$$

where $L$ denotes the lag operator. In this case, the meeting dates only affect the period in which the FOMC meetings occurred so that any changes in volatility as a result of the FOMC meetings do not persist.

In what follows, we will denote the AR(1) specification as SV-AR, the ARMA(1,1) specification as SV-ARMA. Additionally, the short-term interest rate volatility model augmented with an FOMC dummy variable is denoted as SV-AR(MA)X. The next section describes the estimation procedures.

## 2.3 Estimation

The estimation procedures introduced previously are maximum likelihood estimators based on a state space modelling framework. As mentioned in the introduction, state space models can be described by a state equation and a measurement equation. In the context of SV models, the state equation describes the log-variance process whereas the measurement equation describes the observations. For the interest rate model, the observations are the yields and for the crude oil model, the observations are the returns.

We refer readers to Chapter 1, Section 1.1 for details on the description of state space models. In particular, for the QML procedure, the estimation procedure uses the linear state space model given by (1.1) as it uses a transformed linear form of the SV model. On the other hand, the MCL and PF procedures work with the untransformed version of the model and so the state space models are described by (1.2).

Throughout the rest of this chapter, we will use the following notation. We let $y = (y_1, \ldots, y_T)'$ represent the entire vector of observations and $y_{1:t} = (y_1, \ldots, y_t)'$ for $t \leq T$ represent a subset of $y$. Similarly, we let $x = (x_0, \ldots, x_T)'$ represent the entire vector of unobserved state variables and $x_{0:t} = (x_0, \ldots, x_t)'$ represent a subset of $x$.

For most applications related to state space models, the aim is to estimate the latent state process through filtering which involves estimating the marginal distribution of $x_t$ given the observations, i.e. $p(x_t|y_{1:t})$. For the linear Gaussian state space model, the well known Kalman filter can be used to estimate $p(x_t|y_{1:t})$. However, for general SV models including the model given by (2.3) and (2.4), an analytical form for $p(x_t|y_{1:t})$ does not exist. The estimation procedures introduced in the literature consider different approaches to estimate the density.

We will now briefly discuss the approaches taken by each of the three procedures, QML, MCL and PF in the next few sections and demonstrate the practical implementation of the SV models.

### 2.3.1 Quasi-Maximum Likelihood (QML) Estimation

Consider again the residuals defined in Section 2.2, $y_t = \sigma_t r_{t-1}^{\gamma} \epsilon_t$. Then in a manner similar to Ball and Torous (1999), the model can be linearized as follows,

$$\log y_t^2 = \log \sigma_t^2 + 2\gamma \log r_{t-1} + \log \epsilon_t^2, \tag{2.10}$$

where $x_t = \log \sigma_t^2$ for $t = 1, \ldots, T$. If $\epsilon_t \sim N(0, 1)$, then $\log \epsilon_t^2$ is known to follow a $\log \chi^2$ distribution where $\mathbb{E}[\log \epsilon_t^2] = -1.2704$ and $Var[\log \epsilon_t^2] = \pi^2/2$. Under QML

estimation, if it is assumed that $\log \epsilon_t^2 \sim N(-1.2704, \pi^2/2)$[4], then the system of equations is in a linear state space form allowing for estimation using the Kalman filter. Applying the Kalman filter in this manner results in a minimum mean square linear estimator rather than the minimum mean square estimator. However, the assumption of normality allows for parameter estimation to be conducted by maximizing the likelihood, $L(\Theta)$, obtained via the Kalman filter recursions introduced in Section 1.1.

As the parameters of the model can be estimated using the linear Kalman filter, the main advantage of QML is that it is computationally efficient compared to most other proposed methods. Some studies make the comment that QML is sufficient in empirical applications involving SV models, however as will be shown later, the results can vary depending on the estimation procedure used. The standard errors for the parameter estimates are computed using the White (1982) consistent estimate of the asymptotic covariance matrix for the maximum likelihood estimate of the parameter set $\widehat{\Theta}$. This is given by $T^{-1}(\mathfrak{I}_{2D}\mathfrak{I}_{OP}^{-1}\mathfrak{I}_{2D})$ where $\mathfrak{I}_{2D}$ denotes the Hessian estimate of the information matrix and $\mathfrak{I}_{OP}$ denotes the outer product (variance of the score) estimate of the information matrix. These are given by

$$\mathfrak{I}_{2D} = T^{-1}\frac{\partial^2 \log L(\Theta)}{\partial\Theta\partial\Theta'}\Big|_{\Theta=\widehat{\Theta}}, \tag{2.11a}$$

$$\mathfrak{I}_{OP} = T^{-1}\sum_{t=1}^{T}\left(\frac{\partial \log l_t(\Theta)}{\partial\Theta}\Big|_{\Theta=\widehat{\Theta}}\right)\left(\frac{\partial \log l_t(\Theta)}{\partial\Theta}\Big|_{\Theta=\widehat{\Theta}}\right)', \tag{2.11b}$$

where $l_t(\Theta)$ is the marginal likelihood. As these quantities are not known in closed-form, they are evaluated numerically.

We are also able to estimate the posterior mean of the volatility process and produce

[4]It is also possible to leave $Var[\log \epsilon_t^2]$ unspecified and estimate it as a free parameter to improve the performance of QML. However, doing this still resulted in values close to $\pi^2/2$ in the empirical applications. We thus keep it fixed at this value.

forecasts. As $x_t$ above is assumed normal, $\sigma_t^2$ should be lognormal. However, it has been noted by Harvey and Shephard (1993) that due to the non-normality of the SV model after linearization, when $x_t$ is estimated, $x_{t|T}$ does not converge to normality. This means the expectation of $\exp(x_{t|T}) - \exp(x_t)$ is non-zero. Harvey and Shephard (1993) suggest that in practice, the estimate of the variance process should be given by,

$$\mathbb{E}[\sigma_{t|T}^2] = \hat{\sigma}_{t|T}^2 = \exp(\hat{x}_{t|T}), \tag{2.12}$$

where $\hat{x}_{t|T}$ is the smoothed estimate of $\hat{x}_t$ obtained through the Kalman smoother[5]. Similar to the posterior mean estimation, the one-step ahead forecast is constructed as

$$\hat{\sigma}_{T+1|T}^2 = \exp(\hat{x}_{T+1|T}). \tag{2.13}$$

Hence we can determine whether the QML procedure is able to accurately estimate and forecast the variance.

### 2.3.2 Monte Carlo Likelihood (MCL) Estimation

The MCL estimation procedure was first applied to SV models by Sandmann and Koopman (1998) on the linearized model of Harvey et al. (1994) without making the assumption of Gaussian disturbances in the observation equation. This involved estimating Equation (2.10) where $\log \epsilon_t^2$ is estimated as a $\log \chi^2$ random variable using Monte Carlo simulations through an importance sampler. However, as the model was originally specified in a nonlinear form in (2.3) the model parameters are estimated from this form. In the literature, Godambe (1985) and Thavaneswaran

[5]See Harvey (1989) for details about the Kalman smoother.

and Abraham (1988) demonstrate estimating nonlinear models using estimating equations. More recently, Koopman and Lee (2004) considered a nonlinear form of the SV model and in this study (not reported) a comparison of the parameter estimates obtained from either the linear or non-linear versions were found to be similar[6].

The idea of importance sampling under MCL can be explained as follows. If we denote the observations by $y_t$, the latent log-variance by $x_t$ for $t = 1, \cdots, T$ and let $y = (y_1, \cdots, y_T)$ and $x = (x_0, \cdots, x_T)$, then the likelihood function can be expressed as

$$p(y) = \int p(y|x)p(x)dx. \tag{2.14}$$

For the SV model, it is not possible to derive an analytical expression for the integration above due to the multi-dimensional densities involved. One could estimate the above expression using a naïve approach to Monte Carlo simulations as follows,

$$\hat{p}(y) = N^{-1} \sum_{i=1}^{N} p(y|x^{(i)}), \tag{2.15}$$

where $x^{(i)}$ is a draw from $p(x)$. However, this estimator is very inefficient as many of the draws from $p(x)$ would contribute little to $p(y|x)$ and the number of simulations needed to obtain an accurate estimate would be prohibitively large even for relatively small T. Durbin and Koopman (1997) proposed importance sampling techniques to improve the performance of Monte Carlo simulations in the estimation of nonlinear and non-Gaussian state space models. In order to estimate the likelihood, the

[6]The MCL procedure for both the linear and non-linear model was used to estimate the model and resulted in very small differences between the parameter estimates. These differences are most likely due to machine errors as a result of discretization and different Monte Carlo samples used in estimation. As such, the original nonlinear specification was retained in this study.

following result was shown by Durbin and Koopman (1997),

$$\begin{aligned} p(y) &= \int p(y|x)p(x)dx \\ &= \int \frac{p(y|x)p(x)}{q(x|y)} q(x|y)dx \\ &= q(y) \int \frac{p(y|x)p(x)}{q(y|x)q(x)} q(x|y)dx \\ &= q(y)\mathbb{E}_q \left[ \frac{p(y|x)p(x)}{q(y|x)q(x)} \right], \end{aligned} \tag{2.16}$$

where $q(x|y)$ is the importance density and the expectation is with respect to the importance density. The second step above uses Bayes' rule to give

$$q(x|y) = \frac{q(y|x)q(x)}{q(y)}. \tag{2.17}$$

If we consider the stochastic volatility model given by (2.3) and (2.4), then since both $\sigma_t$ and $\epsilon_t$ are stochastic, the model is nonlinear. The methodology relies on approximating the model by a linear Gaussian model where both $q(y|x)$ and $q(x)$ are Gaussian. Furthermore, (2.4) allows us to set $q(x) = p(x)$ as $x_t$ is assumed to be Gaussian. This means the procedure only relies on finding an importance density for $q(x|y)$ in order to perform Monte Carlo simulations. Following Durbin and Koopman (1997) and Koopman and Lee (2004), an approximating Gaussian model is constructed as follows,

$$y_t = a_t + x_t + e_t, \quad e_t \sim \text{NID}(0, b_t), \tag{2.18}$$

where $x_t$ is defined by (2.4). The two parameters, $a_t$ and $b_t$ are estimated to provide a good match of the approximating Gaussian model and the true model. This is achieved by equating the first and second derivatives of the conditional likelihood functions so that the modes of the two distributions coincide. In general the values for $a_t$ and $b_t$ can not be analytically solved as they require knowledge of the value

of the latent state variable. We provide a summary of the iterative procedure to derive the values of $a_t$ and $b_t$ in Appendix 2.A at the end of this chapter. However, for further details see Durbin and Koopman (1997) and Koopman and Lee (2004).

Once the importance density is found, simulations from the importance density can be performed using the simulation smoother of de Jong and Shephard (1995) or Durbin and Koopman (2002). By drawing N sample paths of the state vector $x$, an estimate of the likelihood function is given by the following

$$L(\Theta) = L_G(\Theta)N^{-1}\sum_{i=1}^{N} w(x^{(i)}), \tag{2.19}$$

where $\Theta$ denotes the parameter set, $L_G(\Theta)$ is the Gaussian likelihood based on the importance density $q(x|y,\Theta)$, $x^{(i)}$ is a draw from the importance density $q(x|y,\Theta)$ and

$$w(x^{(i)}) = \frac{p(y|x^{(i)},\Theta)}{q(y|x^{(i)},\Theta)}, \tag{2.20}$$

for $i = 1,\ldots,N$ are the importance weights for each simulated sample path. Since under the SV model we have $y_t|x_t \sim N(0,x_t)$, the two conditional likelihood functions are given by

$$p(y|x) = \prod_{t=1}^{T} \frac{1}{\sqrt{2\pi}r_{t-1}^{\gamma}\exp(x_t/2)} \exp\left(-\frac{y_t^2}{2r_{t-1}^{\gamma}\exp(x_t)}\right), \tag{2.21}$$

$$q(y|x) = \prod_{t=1}^{T} \frac{1}{\sqrt{2\pi}b_t} \exp\left(-\frac{(y_t-a_t)^2}{2b_t^2}\right). \tag{2.22}$$

The likelihood could be interpreted as the Gaussian likelihood of the approximating model scaled by a correction for the departure from Gaussian of the true likelihood. The degree by which the true likelihood deviates from the Gaussian likelihood is governed by the importance weights. In practice, the log-likelihood is maximized

for parameter estimation, however Durbin and Koopman (1997) show that a bias correction is required in this case. They show that an unbiased estimate of the log-likelihood is given by

$$\log L(\Theta) = \log L_G(\Theta) + \log \bar{w} + \frac{s^2}{2N\bar{w}^2}, \tag{2.23}$$

where $\bar{w}$ and $s^2$ are the sample mean and variance of the importance weights respectively. The number of simulations used here is $N = 400$[7] which is large enough so that the bias is quite small and computational time is not too long. When maximizing the likelihood function, the same random numbers should be used in the Monte Carlo simulations for different parameter sets. This ensures the log-likelihood function is smooth with respect to the parameters allowing for the use of gradient-based optimization methods and estimating the standard errors of the parameter estimates. The standard errors in this case are estimated using the numerical Hessian of the log-likelihood based on the maximized likelihood values from one draw of random numbers.

We are also able to obtain the posterior mean and make forecasts of the variance. For the MCL procedure, we can estimate the posterior mean by making use of the draws from the importance density. As the model remains in its original form unlike in the case of the QML procedure, it is appropriate to consider the variance as log-normally distributed. An estimate of the posterior variance is therefore given by

$$\hat{\sigma}^2_{t|T} = \exp(\hat{x}_{t|T} + \frac{1}{2}\hat{P}_{t|T}), \quad t = 1, \ldots, T, \tag{2.24}$$

[7]Simulations are augmented with antithetic variables to increase the number of simulations whilst only marginally increasing computation time.

where both $\hat{x}_{t|T}$ and $\hat{P}_{t|T}$ are estimated from the importance samples as follows,

$$\hat{x}_{t|T} = \sum_{i=1}^{N} w(x^{(i)}) x^i_{t|T}/\bar{w}, \quad \hat{P}_{t|T} = \sum_{i=1}^{N} w(x^{(i)}) P^i_{t|T}/\bar{w}, \quad \bar{w} = \sum_{i=1}^{N} w(x^{(i)}). \tag{2.25}$$

Here $x^{(i)}_{t|T}$ and $P^{(i)}_{t|T}$ correspond to the smoothed state and covariance obtained by applying the Kalman smoother to the Gaussian approximating model for each draw $x^{(i)}$ from the importance density $q(x|y)$. Forecasts of the variance can also be constructed in a similar manner. The one-step ahead forecast of the variance is given by,

$$\hat{\sigma}^2_{T+1|T} = \exp(\hat{x}_{T+1|T} + \frac{1}{2}\hat{P}_{T+1|T}), \quad t = 1, \ldots, T, \tag{2.26}$$

where

$$\hat{x}_{T+1|T} = \sum_{i=1}^{N} w(x^{(i)}) x^i_{T+1|T}/\bar{w}, \quad \hat{P}_{T+1|T} = \sum_{i=1}^{N} w(x^{(i)}) P^i_{T+1|T}/\bar{w},$$

and $\bar{w}$ is the same as for the smoothed estimates. In this case, $x^i_{T+1|T}$ and $P^i_{T+1|T}$ are obtained from the Kalman filter applied to the Gaussian approximating model for each draw $x^{(i)}$ from the importance density $q(x|y)$.

### 2.3.3 Particle Filter (PF) Estimation

Particle filtering techniques (also known as Sequential Monte Carlo techniques) have been applied in a number of different disciplines but are becoming increasingly popular in financial econometrics. They were developed in order to estimate and sample from the posterior distribution of nonlinear and/or non-Gaussian state space models when it is not possible to sample directly from the posterior density. Kim et al. (1998) and Pitt and Shephard (1999) were the first authors to consider particle

filtering methods for SV models, although mainly for filtering purposes rather than parameter estimation. However, Pitt (2002) introduced a smooth particle filter which allows for maximum likelihood estimation of the parameters in SV models. As the procedure uses importance sampling, the technique is similar to MCL except that the procedure is a sequential filter where samples are drawn iteratively from an importance density. On the other hand, MCL uses a simulation smoother to generate draws from an importance density. A further difference is that whilst MCL requires a Gaussian importance density due to the use of the Kalman filter and smoother in determining the approximating model, the only restriction for the particle filter is that the importance density can be simulated from. Whilst this allows for more flexibility in the applications available under the particle filter, this usually comes at the cost of efficiency in estimation as an optimal importance density is generally either not available analytically nor computationally efficient to use in Monte Carlo simulations. This has resulted in a number of variations to the particle filter to improve its efficiency, although most of the recent incarnations have evolved from the 'bootstrap filter' approach described in Gordon et al. (1993). For interested readers, an overview of recent advances in particle filtering techniques that have followed Gordon et al. (1993) can be found in Doucet et al. (2001) and Cappé et al. (2007). However, in this chapter, we consider the smooth particle filter of Pitt (2002) which allows for maximum likelihood estimation of SV models.

In order to perform parameter estimation using particle filtering methods, we need to evaluate the likelihood function given the parameters, $L(\Theta) = p(y|\Theta)$ where $\Theta$ denotes the parameter set. However, when using particle filters that incorporate a resampling step, the likelihood function will not be smooth with respect to the parameter set, even if the same random numbers are drawn for each run of the particle filter. This happens because the resampling step essentially involves sampling from the discrete cdf generated by the normalized importance weights. When the

parameters are changed, the importance weights will have changed resulting in a different resampling distribution. If even only a few particles resampled during the early time steps are different, it is clear that this can result in very different particle values as the different particles will be propagated through all later time steps. This could lead to quite different likelihood values even when the parameter values have only changed a little. Pitt (2002) proposed a solution that incorporates a smooth bootstrapping procedure when resampling. This results in a likelihood function that is smooth with respect to the parameter set. Having a smooth likelihood function is desirable for a number of reasons including allowing the use of gradient-based methods for maximizing the likelihood function and computing standard errors for the parameter estimates. The main weakness of this approach is that it can only be applied to one-dimensional state variables strictly limiting its application, although it is suitable for SV modelling application studied here.

In what follows, it is assumed that the likelihood is dependent on the parameter set but we drop $\Theta$ in the notation for simplicity. If we consider the state space model defined by (1.2), then Doucet et al. (2000) show that the likelihood can be decomposed as follows,

$$p(y) = p(y_1) \prod_{t=2}^{T} p(y_t|y_{1:t-1}), \tag{2.27}$$

where $p(y_t|y_{1:t-1})$ is the conditional likelihood or alternatively the log-likelihood is given by

$$\log p(y_{1:T}) = \log p(y_1) + \sum_{t=2}^{T} \log p(y_t|y_{1:t-1}).$$

It is shown in Pitt (2002) that when the importance density used for simulation is assumed to satisfy the Markov property, then the conditional likelihoods can be

estimated by

$$\hat{p}(y_t|y_{1:t-1}) = \mathbb{E}[w_t(x_{0:t})], \tag{2.28}$$

where $w_t(x_{0:t})$ are the importance weights defined by

$$w_t(x_{0:t}) = \frac{p(y_{1:t}|x_{0:t})p(x_{0:t})}{q(x_{0:t}|y_{1:t})}, \tag{2.29}$$

and $q(x_{0:t}|x_{0:t-1}, y_{1:t})$ is the importance density. As shown in Doucet et al. (2000) and Cappé et al. (2007) among others, if it is assumed that the distribution of the states are Markov and that the observations are conditionally independent given the states, then the importance weights can be represented recursively as,

$$w_t = w_{t-1}\frac{p(y_t|x_t)p(x_t|x_{t-1})}{q(x_t|x_{0:t-1}, y_{1:t})}. \tag{2.30}$$

By sampling from the importance density at each step, Pitt (2002) shows that an unbiased estimate of the conditional likelihoods is given by

$$\hat{p}(y_t|y_{1:t-1}) = \frac{1}{N}\sum_{i=1}^{N} w_t^{(i)}, \tag{2.31}$$

where $w_t^{(i)}$ are the importance weights associated with the $i^{\text{th}}$ draw from the importance density. One final consideration is the choice of importance density. In general and in the case of SV models, an optimal importance density does not exist. A popular alternative choice of importance density is the prior density, i.e. $q(x_t|x_{0:t-1}, y_{1:t}) = p(x_t|x_{t-1})$. If this choice of importance density is substituted into (2.30), the weights are then given by the following simple representation,

$$w_t^{(i)} = w_{t-1}^{(i)} p(y_t|x_t^{(i)}), \tag{2.32}$$

where for the interest rate SV model,

$$p(y_t|x_t) = \frac{1}{\sqrt{2\pi} r_{t-1}^{\gamma} \exp(x_t/2)} \exp\left(\frac{y_t^2}{2r_{t-1}^{2\gamma} \exp(x_t)}\right). \tag{2.33}$$

The main drawback of this choice of proposal distribution is that the most recent observation is not considered when generating the particles. With the addition of the resampling step, this is mitigated substantially and is not considered much of an issue. However, as mentioned previously the resampling step causes a non-smooth likelihood function.

Pitt (2002) addressed this issue by proposing a modification to the resampling procedure. It involves firstly sorting the sampled particles and constructing an empirical cdf for the particles at each time step. During resampling, rather than resampling directly from the discrete empirical cdf generated by the importance weights, a smooth and continuous cdf is constructed by linearly interpolating between each particle in the discrete cdf. The particles are resampled from the smooth cdf resulting in resampled particles that are close for small changes in the parameters. An algorithm for the smooth resampling procedure can be found in Pitt (2002) and to conserve space it is not reproduced here. It should also be noted that the construction and resampling from the smooth cdf results in some additional computational cost over SIR as a result of sorting the particles at each time step. As sorting can be completed in O(NlogN) time[8], the complexity of each run of the smooth particle filter is O(TNlogN) and is found here and by Pitt (2002) to increase computational time only slightly. Going by the number of particles typically used in the literature for PF techniques and to keep a balance between computation time and accuracy, the number of particles used here is set to $N = 2000$. A brief algorithm of the smooth particle filter used in this chapter can be found in the Appendix 2.B but

[8]Commonly known sorting algorithms with O(NlogN) complexity include merge sort, heap sort or quick sort. In this thesis, quick sort is used.

more interested readers should refer to Pitt (2002).

As we evaluate the log-likelihood in practice rather than the likelihood itself this introduces a bias due to the Monte Carlo sampling error of the weights similar to the MCL procedure. Pitt (2002) gives an unbiased estimate of the log-likelihood function as the sum of the marginal log-likelihoods where the marginal log-likelihood, $l_t(\theta)$, is given by

$$\log l_t(\Theta) = \log \sum_{i=1}^{N} w_t^{(i)} - \frac{s^2}{2N\bar{w}^2}, \tag{2.34}$$

where $\bar{w}$ and $s$ are the sample mean and standard deviation of the unnormalized importance weights, $w_t^{(i)}, i = 1, \ldots, N$. Hence the log-likelihood is estimated as

$$\log L(\Theta) = \log l_1(\Theta) + \sum_{t=2}^{T} \log l_t(\Theta), \tag{2.35}$$

which is maximized for parameter estimation. To estimate the standard errors, Pitt (2002) uses the variance of the scores (outer-product estimator) as an estimate of the variance-covariance matrix and we follow the same approach.

Having estimated the parameters, we can estimate the posterior mean and make forecasts conditionally on the parameters in a similar manner as MCL. To estimate the posterior mean, we require smoothed estimates of $x_t$ which can be obtained by applying smoothing techniques based on particle filters. Smoothing techniques in particle filtering are generally computationally expensive procedures but as we only need to run them once after parameter estimation, this does not pose a problem here. We consider here the smoothing procedure of Doucet et al. (2000) which involves estimating the posterior density $p(x_{t|T})$ by Monte Carlo simulation. They

approximate the posterior density using the following expression,

$$\hat{p}(dx_t|y_{1:T}) = \sum_{i=1}^{N} \tilde{w}_{t|T}^{(i)} \delta_{x_t^{(i)}}(dx_t), \tag{2.36}$$

where $\tilde{w}_{t|T}^{(i)}$ are the normalized smoothing weights and $\delta_{x_t^{(i)}}(\cdot)$ is the Dirac delta function. This means the posterior density has the same draws ($x_t^{(i)}$, $i = 1, \ldots, N$, $t = 1, \ldots, T$) as the filtering density but with different weights. The smoothing weights are obtained recursively as follows,

$$\tilde{w}_{t|T}^{(i)} = \sum_{j=1}^{N} \tilde{w}_{t+1|T}^{(j)} \frac{\tilde{w}_t^{(i)} p(x_{t+1}^{(j)}|x_t^{(i)})}{\sum_{k=1}^{N} \tilde{w}_t^{(j)} p(x_{t+1}^{(j)}|x_t^{(k)})}, \tag{2.37}$$

where $\tilde{w}_t^{(i)} = w_t^{(i)} / \left(\sum_{j=1}^{N} w_t^{(i)}\right)$ are the normalized importance weights and $w_t^{(i)}$ comes from Equation (2.32). The smoothed estimates of the posterior mean and covariance of $x_t$ are given by,

$$\hat{x}_{t|T} = \sum_{i=1}^{N} \tilde{w}_{t|T}^{(i)} x_t^{(i)}, \quad \hat{P}_{t|T} = \sum_{i=1}^{N} \tilde{w}_{t|T}^{(i)} (x_t^{(i)} - \hat{x}_{t|T})^2. \tag{2.38}$$

An estimate of the posterior mean of the smoothed variance $\hat{\sigma}_{t|T}$ is given by Equation (2.25). To make conditional forecasts of the state variables under the PF procedure, Doucet et al. (2000) suggest the following approach. The one-step ahead forecast for the state involves firstly sampling from the conditional density $p(x_{T+1}|x_T^{(i)})$, $i = 1, \ldots, N$ to obtain the random samples, $x_{T+1}^{(i)}$. An estimate of the conditional density is given by

$$\hat{p}(dx_{T+1}|y_{1:T}) = \sum_{i=1}^{N} \tilde{w}_T^{(i)} \delta_{x_{T+1}^{(i)}}(dx_{T+1}). \tag{2.39}$$

Hence the forecast mean and covariance can be estimated by

$$\hat{x}_{T+1|T} = \sum_{i=1}^{N} \tilde{w}_T^{(i)} x_{T+1}^{(i)}, \quad \hat{P}_{T+1|T} = \sum_{i=1}^{N} \tilde{w}_T^{(i)} (x_{T+1}^{(i)} - \hat{x}_{T+1|T})^2. \tag{2.40}$$

The one-step ahead forecast of the variance, $\sigma^2_{T+1}$ can be obtained from (2.26). Under both the PF and MCL procedure, the smoothing and forecasting of the states is fairly straightforward requiring only a little additional computational effort.

### 2.3.4 Implementation

In terms of ease of implementation, QML was the simplest as it only requires application of the Kalman filter. Both MCL and PF were fairly straightforward but required significantly more effort to implement. For the MCL, this was a two stage process where in the first step, the importance is derived, and in the second step the simulations were conducted from the importance density. For the particle filter, although the general algorithm was fairly straightforward, the most difficulty lay in the smoothing algorithm. The estimation procedures were all implemented using Matlab and C[9]. As a comparison of computational efficiency, we report the times to evaluate the likelihood function for each procedure. The times are based on the average time to evaluate the likelihood function for a time-series of length $T = 1000$ using the standard AR(1) SV model on a 2.53Ghz processor. The QML takes 0.04s, MCL takes 3.75s and PF takes 4.91s.

### 2.3.5 Finite Sample Performance

Although a number of studies have compared the finite sample performance of QML and MCL in the literature, PF has not received much attention in terms of parameter

[9]Matlab was used in conjunction with C to improve computational efficiency through the use of Mex files. Matlab is known to handle loops very inefficiently and for the recursive steps in each estimation procedure, C was used instead. This makes no negligible difference to the results as using either language for the loops results in differences close to double precision (i.e. 1e-16) whilst substantially reducing computation time.

estimation. In this section, a brief study of the finite sample performance of each of the three procedures is conducted using the standard AR(1) model. In this case, we consider the model given by (2.3) - (2.4) with $\gamma = 0$ and where $y_t$ represents the demeaned returns. To conduct the study, 500 sample paths of the volatility and return process are simulated using the same parameter set. The length of each sample path is set to $T = 1000$ where it is assumed that the frequency of returns is weekly. For the traditional AR(1) model, the parameter set is $\Theta = (\phi, \sigma_\eta, \mu)$. In this application we set the parameters to that commonly found in the literature on empirical studies of equity returns. In some of these studies such as Jacquier et al. (1994) or Sandmann and Koopman (1998), a typical parameter set found for S&P500 returns is $\Theta = (0.96, 0.2, 1)$.

**Table 2.1**
Simulation Results

This table reports the mean of the parameter estimates for each estimation procedure based on 500 simulations of a return series with stochastic volatility with sample paths of length $T = 1000$. The row labelled 'True' denotes the actual parameter set used in the simulations. The RMSE of the estimates are in parentheses.

| | Parameter | | |
|---|---|---|---|
| | $\phi$ | $\sigma_\eta$ | $\mu$ |
| True | 0.9500 | 0.2000 | 1.0000 |
| QML | 0.9434 | 0.2244 | 1.0100 |
| | (0.0422) | (0.0916) | (0.1740) |
| MCL | 0.9519 | 0.2081 | 1.0090 |
| | (0.0221) | (0.0427) | (0.1675) |
| PF | 0.9520 | 0.2098 | 1.0147 |
| | (0.0219) | (0.0432) | (0.1715) |

The simulation results can be found in Table 2.1. The QML and MCL results tend to agree with previous findings in the literature with slightly higher bias exhibited by QML in the $\phi$ and $\sigma_\eta$ estimates. Interestingly, although the two importance sampling procedures take fundamentally different approaches, PF shows very similar performance to MCL. This is true for both the mean of the parameter estimates

and the root mean-squared errors (RMSE) of the estimates. For both the MCL and PF procedures, the RMSE of $\phi$ and $\sigma_\eta$ is around half the size of the RMSE for the QML estimates. This demonstrates that PF is able to correct for some of the finite sample bias exhibited by QML.

## 2.4 Empirical Results

This section conducts an empirical study using short-term interest rate and crude oil data. The models outlined in Section 2.2 are estimated using the three estimation procedures for each data set. We firstly provide some descriptive statistics of the data and compare the parameter estimates obtained under the various models and estimation procedures. We then evaluate the performance of the models and look at whether the choice of model is affected by the estimation procedure. Finally, we conduct a comparison of the performance of each model and estimation procedure when conducting out-of-sample volatility forecasts.

### 2.4.1 Data

The short-term interest rate data consisting of the yield on three-month constant maturity U.S. treasury bills and the FOMC meeting dates are obtained from the Board of Governors of the Federal Reserve website[10]. The crude oil spot price data is obtained from the Energy Information Administration (EIA) website[11]. Both the short-term interest rate and crude oil data consists of weekly observations from January 1990 to May 2006 giving a sample size of 856 observations for each data set.

[10] http://www.federalreserve.gov.
[11] http://www.eia.doe.gov.

The estimation procedure used for this chapter follows the approaches taken previously in the literature where first OLS is performed to estimate $a$ and $b$ in Equation (2.2) for interest rates, or Equation (2.6) for crude oil. Once this is estimated, the residuals from the regression are used in order to estimate the volatility process $\sigma_t$ where for interest rates $y_t = \sigma_t r_{t-1}^{\gamma} \epsilon_t$ and for crude oil $y_t = \sigma_t \epsilon_t$. Descriptive statistics for the interest rate and crude oil data can be found in Table 2.2 including the residuals from the CKLS discretization. The plots of the data can be found in Figure 2.1.

**Table 2.2**
Description of Data

Panel A reports the linear regression estimates defined by (2.2) for short-term interest rates and (2.6) for crude oil. Panel B reports some descriptive statistics for weekly observations of short-term interest rates (SR) and the crude oil (CO) data for the period January 1990 to May 2006. JB is the Jarque-Bera test for normality, Q(20) is the Ljung-Box portmanteau test for serial correlation at a lag of 20, ADF is the augmented Dickey-Fuller test for unit root and ARCH LM is the autoregressive conditional heteroscedasticity Lagrange multiplier test at a lag of 20. P-values for the test statistics are in parentheses.

Panel A: Linear Regression

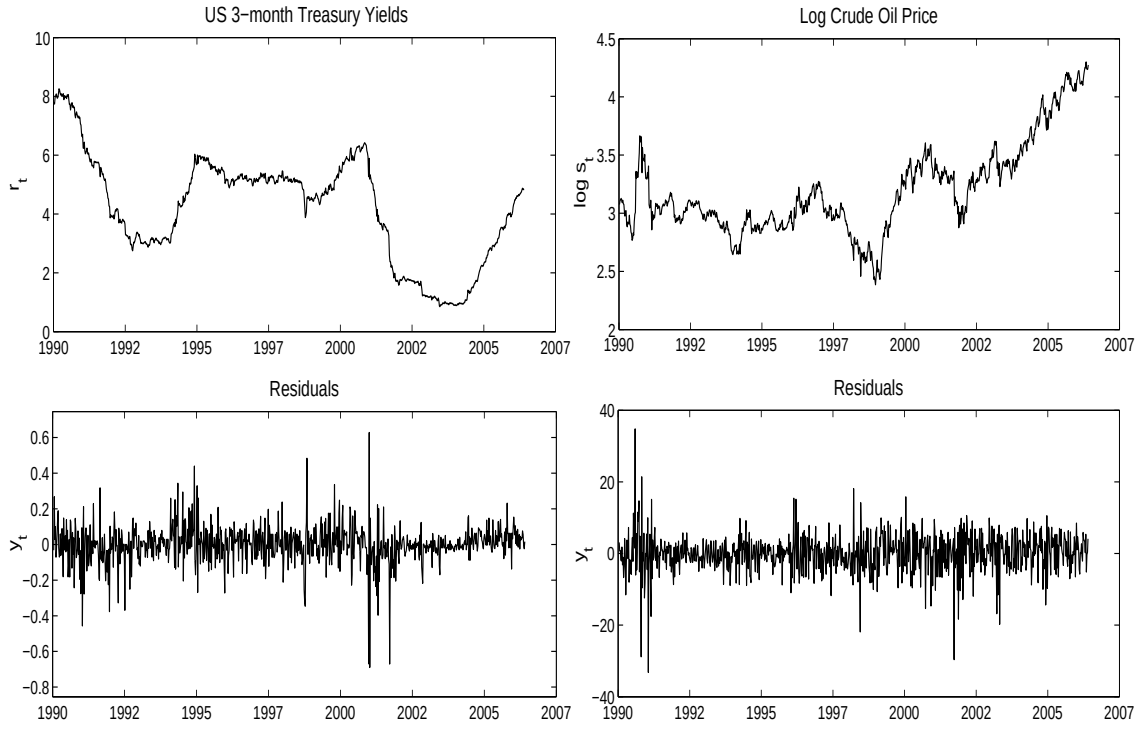
| | SR | | CO | |
|---|---|---|---|---|
| Parameter | $a$ | $b$ | $a$ | $b$ |
| Estimate | 0.0001 | −0.0040 | 1.7701 | −0.0051 |
| (S.E.) | (0.0000) | (0.0020) | (1.5213) | (0.0047) |

Panel B: Descriptive Statistics

| | Mean | S.D. | Skew. | Kurt. | JB | Q(20) | $Q^2(20)$ | ADF | ARCHLM |
|---|---|---|---|---|---|---|---|---|---|
| SR | | | | | | | | | |
| $r_t$ | 4.187 | 1.808 | -0.156 | 2.398 | 16.39 | 26270.63 | 86.46 | -1.64 | 844.67 |
| | | | | | (0.000) | (0.000) | (0.000) | (0.095) | (0.000) |
| $y_t$ | 0.000 | 0.108 | -0.680 | 10.889 | 2283 | 166.16 | 187.70 | -31.33 | 127.12 |
| | | | | | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| CO | | | | | | | | | |
| $\log s_t$ | 3.195 | 0.391 | 0.862 | 3.366 | 110.884 | 13628.72 | 13663.83 | 0.626 | 822.55 |
| | | | | | (0.000) | (0.000) | (0.000) | (0.851) | (0.000) |
| $y_t$ | 0.000 | 5.377 | -0.540 | 9.053 | 1332.93 | 49.46 | 173.65 | -33.45 | 106.12 |
| | | | | | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |

If we firstly consider the estimated coefficients of the linear regressions in the top panel of Table 2.2, the estimate of b is negative for both the interest rate and crude oil data indicating that they are mean-reverting. However, the mean-reversion rates are very low and the mean-reversion of the crude oil data is not significant. This

can be seen in Figure 2.1b where both the yield and log crude oil price plots show that the plots stay within a band of values most of the time, although they also exhibit some periods of steep increases and decreases. Figure 2.1 also shows the residuals from the regression where evidence of both time-varying volatility and volatility clustering can be observed in the plots of $y_t$. This can be observed by periods of high volatility and low volatility occuring in "clusters" for both data sets. Two such periods of high volatility in both the yields and crude oil prices can be identified during the period surrounding the Iraq war around 1990-92 and the period immediately following September 11, 2001.




**(a)** Interest Rate Data **(b)** Crude Oil Data

**Figure 2.1**
Plots of yields, crude oil prices and regression residuals, January 1990 - May 2006. The yields and residuals are in percentage points. The log crude oil price is in $US.

Looking at the descriptive statistics of the data in the bottom panel of Table 2.2 both the yield and crude oil data exhibit some similar characteristics. It can be seen from the augmented Dickey-Fuller (ADF) unit root test statistic that non-stationarity can not be rejected for the yields and crude oil prices which is consistent with what was

observed in regards to the mean-reversion rates. The residuals of both series exhibit negative skewness, high kurtosis and the Jarque-Bera statistic indicates strong non-normality. There is some evidence of autocorrelation in both residuals series as indicated by the Ljung-Box (LB) Q test statistic, although of particular interest is autocorrelation squared residuals. Applying the LB test to the squared residuals and testing for autoregressive conditional heteroscedasticity (ARCH) effects gives an indication of whether time-varying volatility is present. The results of both tests indicate that the residuals exhibit autocorrelation in the squared residuals and ARCH effects for both the yield data and crude oil data. This suggests that volatility is time-varying in both data series and that a stochastic volatility model may be appropriate.

### 2.4.2 Parameter Estimates

The results of the maximum likelihood parameter estimate can be found in the top panel of Tables 2.3 and 2.4 for the interest rate and crude oil data respectively. Each table lists the parameter estimates and their standard errors under each of the different estimation techniques described in Section 2.3.

#### Short-Term Interest Rates

If we examine the values of the parameter estimates, it is clear from the tables that the parameter estimates of the three estimation procedures differ by varying degrees even though most parameters for all models are statistically significant at the 5% level. The parameter estimates for the short-term interest rate data can be found in Table 2.3. We firstly concentrate on the estimates of the basic SV-AR specification.

The main difference in parameter estimates for the three procedures can be found in the two main parameters of the model, $\phi$ and $\sigma_\eta$, commonly referred to as the persistence in volatility and the volatility of volatility respectively. Whilst $\phi$ governs how quickly the models revert to the long-term mean level of volatility, $\sigma_\eta$ could be considered as the main driver in the stochastic nature of volatility. Differences in both these values can lead to significant differences in the conditional variances estimated. It has been noted in Ruiz (1994) that as $\phi$ and $\sigma_\eta$ decrease, the finite sample bias of the QML estimates increases for these parameters. This seems to be the case here where the QML estimates of $\phi$ and $\sigma_\eta$ were found to be approximately 0.58 and 0.95 respectively which tend to differ from the MCL estimates of 0.68 and 0.75 and the PF estimates of 0.65 and 0.77. For this sample, it seems that QML underestimates $\phi$ and as a result, overestimates $\sigma_\eta$. This agrees with the finite sample performance of QML estimates observed previously in the literature (see e.g. Jacquier et al. (1994), Sandmann and Koopman (1998) and Broto and Ruiz (2004)) and the brief study above. For $\mu$ and $\gamma$, both are estimated very closely between MCL and PF, whilst again a fairly substantial difference is evident with the QML estimate. The value of $\gamma$ is around 0.59 for both MCL and PF whereas it is higher for QML at 0.65. These are quite close to a value of 0.5 which implies that the CIR model of interest rates holds when stochastic volatility is incorporated. This is consistent with Ball and Torous (1999) and provides further evidence that the original value of 1.5 estimated in CKLS is too high.

Similar results are exhibited for the other specifications when comparing the estimation procedures. Looking at the ARMA estimates in Table 2.3, the QML estimate of $\theta$ is negative and not significant. The estimate for $\theta$ under both MCL and PF is significant at the 5% level with a value close to 1 although the standard errors are quite high. This implies that the level of volatility is highly correlated with past shocks or innovations. However, the inclusion of the MA term lowers both the $\phi$

**Table 2.3**
Parameter Estimates for Short-Term Interest Rates

This table provides the parameter estimates for each model specification and grouped by estimation procedure for the short-term interest rate data. Standard errors for the parameter estimates are in parentheses.

| | Parameter | | | | | |
|---|---|---|---|---|---|---|
| Model | $\phi$ | $\theta$ | $\sigma_\eta$ | $\mu$ | $\gamma$ | $\kappa$ |
| QML | | | | | | |
| SV-AR | 0.5783 | – | 0.9511 | −6.9195 | 0.6569 | – |
| | (0.1090) | – | (0.1584) | (0.2698) | (0.0950) | – |
| SV-ARMA | 0.6402 | −0.1950 | 1.0332 | −6.9194 | 0.6569 | – |
| | (0.1338) | (0.2725) | (0.1760) | (0.2732) | (0.0950) | – |
| SV-ARX | 0.5756 | – | 0.9463 | −6.9847 | 0.6561 | 0.4238 |
| | (0.1101) | – | (0.1584) | (0.2698) | (0.0945) | (0.2234) |
| SV-ARMAX | 0.6320 | −0.1779 | 1.0403 | −6.9843 | 0.6560 | 0.4223 |
| | (0.1377) | (0.2827) | (0.1790) | (0.2734) | (0.0955) | (0.2252) |
| MCL | | | | | | |
| SV-AR | 0.6238 | – | 0.7492 | −6.6990 | 0.5884 | – |
| | (0.0701) | – | (0.0740) | (0.2209) | (0.0777) | – |
| SV-ARMA | 0.5096 | 0.9913 | 0.4776 | −6.6907 | 0.5880 | – |
| | (0.1062) | (0.3866) | (0.0681) | (0.2188) | (0.0769) | – |
| SV-ARX | 0.6278 | – | 0.7194 | −6.7835 | 0.5894 | 0.5394 |
| | (0.0734) | – | (0.0768) | (0.2194) | (0.0766) | (0.1684) |
| SV-ARMAX | 0.5359 | 0.9943 | 0.4597 | −6.7806 | 0.5903 | 0.5402 |
| | (0.1122) | (0.3634) | (0.0529) | (0.2149) | (0.0750) | (0.1654) |
| PF | | | | | | |
| SV-AR | 0.6538 | – | 0.7677 | −6.7412 | 0.5943 | – |
| | (0.0590) | – | (0.0575) | (0.2384) | (0.0858) | – |
| SV-ARMA | 0.5655 | 0.9991 | 0.4663 | −6.7176 | 0.5930 | – |
| | (0.0924) | (0.4701) | (0.0685) | (0.2304) | (0.0823) | – |
| SV-ARX | 0.6565 | – | 0.7365 | −6.8119 | 0.5932 | 0.5288 |
| | (0.0734) | – | (0.0843) | (0.2383) | (0.0840) | (0.1535) |
| SV-ARMAX | 0.5823 | 0.9991 | 0.4355 | −6.8155 | 0.5989 | 0.5255 |
| | (0.0919) | (0.1291) | (0.0718) | (0.2274) | (0.0812) | (0.1469) |

and $\sigma_\eta$ values compared to the AR(1) model. Of some concern is the MA value close to 1 for both the MCL and PF estimates as it implies that the processes are close to non-invertible. However, since a non-invertible MA process is still a stationary process, this does not pose a problem for the model specification and the results remain valid.

The inclusion of the FOMC dummy to the SV model has less of an effect than an MA term on the other parameter estimates as can be seen in Table 2.3. Moreover, $\kappa$ is significant at the 5% level for both MCL and PF, but only significant at the 10% level for the QML estimate. Whilst the parameter estimate differs for all three estimation procedures, it is positive for all three which indicates that FOMC meetings increases the volatility of interest rates. Again there appears to be some differences between the QML estimates of around 0.42 compared to the MCL and PF estimates of around 0.53-0.54. However, it should be noted that the estimates imply that during the week surrounding FOMC meetings, volatility increased by around 20% under QML, and around 30% under MCL and PF. There is quite a substantial increase in volatility in all three cases. This is consistent with the findings in Das (2002) where the volatility of the federal funds rate increased when meetings of the FOMC occur. Das (2002) considered a jump-diffusion model of the federal funds rate and found that the frequency of jumps increased after two-day FOMC meetings consistent with an increase in volatility. There are two possible explanations for the increase in volatility. It could suggest that either markets are efficient so that trading activity and volatility increases when new information arrives or that surprise announcements after the meetings substantially increase volatility of interest rates. Whilst it is difficult to distinguish between the two possibilities without looking at additional data, the large increase in volatility implied by the estimates suggests that surprise announcements during the meetings have a large impact on volatility. It should also be noted that all parameter estimates except $\kappa$ are rela-

tively unaffected compared to the models without the FOMC factor. This means that the impact of FOMC meetings on interest rate volatility is not captured by the other parameters in the model. Although this only applies to this sample, this suggests that when modelling volatility in bond markets, the activities of central banks should be accounted for.

**Crude Oil**

The results of the crude oil parameter estimates can be found in Table 2.4 where all parameters are significant at the 5% level for each model and estimation procedure. The parameter estimates are substantially different to the short-term interest rate results indicating that the dynamics of volatility for both markets are largely different. In fact, the parameter estimates tend to be similar to those reported for equity markets which is not surprising given the parameterisation of the model is closer to those typically applied in equity and foreign exchange market studies. In terms of the differences between estimation procedures, the results are slightly different to the interest rate results. Concentrating on the SV-AR specification first, the persistence parameter, $\phi$, is estimated slightly higher under QML a value of 0.967 and 0.958 under both MCL and PF. On the other hand, the QML estimate for $\sigma_\eta$ of 0.161 is lower than the MCL and PF estimates of 0.229 and 0.231 respectively. Although these differences are lower than for the interest rate results, this is the opposite result when comparing the QML estimates to MCL and PF. In this case, QML overestimates $\phi$ and underestimates $\sigma_\eta$ although the result is not contradictory to previous studies since this holds only for this sample. Also, it seems that the difference between the QML estimate of $\phi$ to the MCL and PF estimates is much smaller than the difference in the estimate of $\sigma_\eta$. This is consistent with the finding in Ruiz (1994) who found a larger bias when either $\phi$ or $\sigma_\eta$ is low with the bias

decreasing when these two parameters increase. Finally, there is little difference between the long-term mean volatility parameter, $\mu$, similar to the interest rate results.

Similar results hold for the SV-ARMA specification. There appears to be some differences in the QML estimates of $\phi$, $\theta$ and $\sigma_\eta$ compared to MCL and PF estimates although they are not large. Both $\phi$ and $\mu$ differ only slightly from the SV-AR values and as with the interest rate data, $\sigma_\eta$ differs somewhat from the SV-AR values. However, contrary to the interest rate data, $\sigma_\eta$ is higher under the SV-ARMA specification and $\theta$ is negative and quite high for all procedures. This suggests that volatility is negatively correlated with past shocks in volatility and that an increase in volatility in the previous period will likely be followed by a decrease in volatility in the following period.

Whilst the evidence is not conclusive, it is clear from the two data sets that the parameter estimates under QML differ from either MCL and PF more than the MCL and PF estimates differ. When comparing the MCL and PF estimates, the only differences arise from the estimates of $\phi$ and $\sigma_\eta$ although they are relatively small compared to the difference in the QML estimates. It is also interesting to note that when using the smooth version of the particle filter, the estimated standard errors are quite close to the standard errors estimated under MCL for both data sets. This is encouraging for the smooth PF estimation procedure, as whilst it is noted in Pitt (2002) that the standard errors estimated for linear Gaussian state space models are close to the values estimated under the Kalman filter, no comparison was made for nonlinear or non-Gaussian state space models. Although this result holds only for these particular data sets and models, this shows promise for both the MCL and PF procedures in other applications involving nonlinear and/or non-Gaussian models.

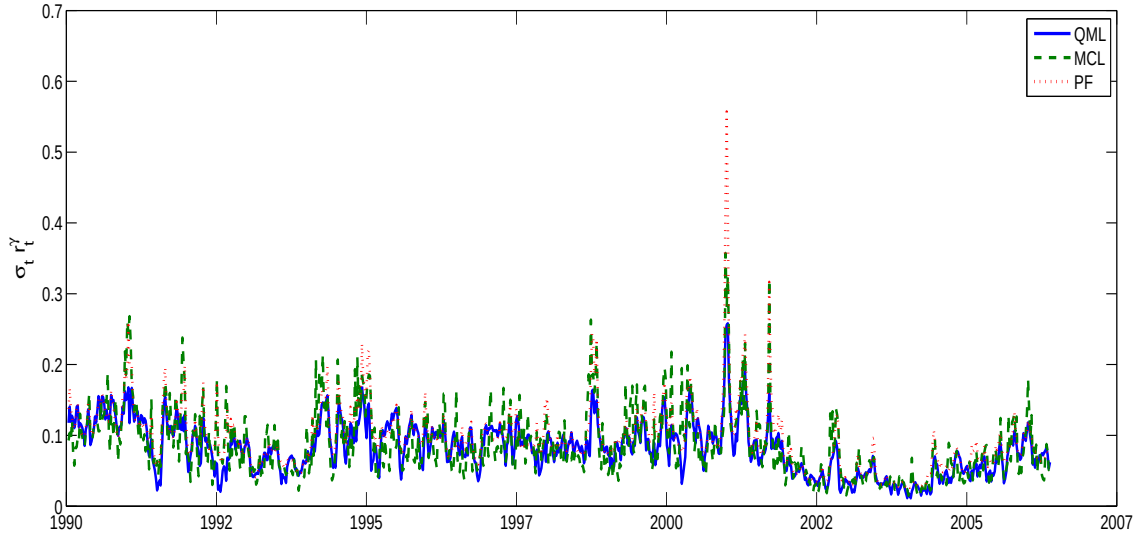
**Table 2.4**
Parameter Estimates for Crude Oil

This table provides the parameter estimates for each model specification and grouped by estimation procedure for the crude oil data. Standard errors for the parameter estimates are in parentheses.

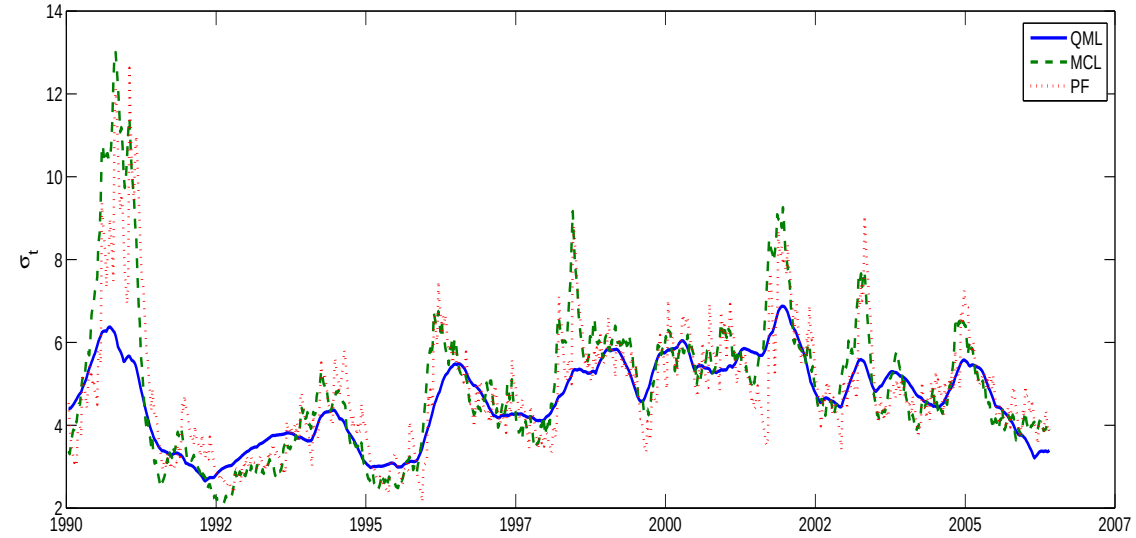
| | Parameter | | | |
|---|---|---|---|---|
| Model | $\phi$ | $\theta$ | $\sigma_\eta$ | $\mu$ |
| QML | | | | |
| SV-AR | 0.9666 | – | 0.1608 | 2.9823 |
| | (0.0794) | – | (0.0578) | (0.1780) |
| SV-ARMA | 0.9709 | −0.7799 | 0.6566 | 2.9828 |
| | (0.0787) | (0.1868) | (0.2189) | (0.1823) |
| MCL | | | | |
| SV-AR | 0.9580 | – | 0.2292 | 2.9963 |
| | (0.0557) | – | (0.0421) | (0.1903) |
| SV-ARMA | 0.9670 | −0.6251 | 0.5045 | 2.9473 |
| | (0.0528) | (0.0612) | (0.0902) | (0.1999) |
| PF | | | | |
| SV-AR | 0.9584 | – | 0.2319 | 3.0319 |
| | (0.0521) | – | (0.0371) | (0.1279) |
| SV-ARMA | 0.9670 | −0.6406 | 0.5241 | 2.9903 |
| | (0.0552) | (0.1443) | (0.0850) | (0.2085) |

A comparison of the estimated volatility can also be made from the smoothed estimates of each of the filters. Plots of the volatility estimated under all three estimation procedures for the SV-AR(1) model are shown in Figure 2.2 for both the interest rate and crude oil data. The plots show that whilst the volatility estimates for the interest rate data tends to be fairly similar, the crude oil volatility estimates tend to differ between the estimation procedures. The QML volatility estimates vary less than either the MCL and PF volatility estimates, although the PF estimates appear to be the most volatile. This is most evident during periods of high volatility including during the Gulf War of 1990-1991 for the crude oil data and around September 11, 2001 for both sets of data. In either case, the QML estimates tend to be much lower than the volatility estimates of either MCL or PF. In the

next section we will determine how well volatility is estimated using a number of diagnostic tests.




**(a)** Interest Rate Volatility




**(b)** Crude Oil Volatility

**Figure 2.2**
Plots of volatility estimated under the AR(1) model for each estimation procedure

## 2.5 Model Tests

### 2.5.1 In-Sample Tests

Having compared the parameter estimates using the different estimation procedures, we now consider the difference in model selection criteria. The first set of criteria to test model specification are the likelihood ratio tests, the Akaike Information Criterion (AIC) and Bayesian (Schwartz) Information Criterion (BIC). For the likelihood ratio tests, the AR model was taken as the base model and tested against each of the other specifications. For instance, testing against the ARMA specification, the null hypothesis is $H_0 : \theta = 0$ and the alternative hypothesis is $H_a : \theta \neq 0$. For the AIC and BIC values, the model is chosen according to the lowest value obtained under the two measures where the models are penalized depending on the number of parameters to be estimated.

All three procedures also allow for estimation of the latent volatility process as described in Section 2.3. We are thus able to determine the estimation performance with goodness-of-fit tests. The first metrics we apply to evaluate the goodness-of-fit are the root mean squared error (RMSE), mean absolute error (MAE) and Theil's inequality coefficient (TIC). Similar to Smith (2002), Kalimipali and Susmel (2004) and Sun (2005), we take the absolute residuals as a proxy for volatility and compute the RMSE and TIC as follows,

$$\text{RMSE} = \sqrt{1/T \sum_{t=1}^{T} \Big( |y_t| - \hat{\sigma}_{t|T} r_{t-1}^{\gamma} \Big)^2}, \tag{2.41}$$

$$\text{MAE} = 1/T \sum_{t=1}^{T} \Big| |y_t| - \hat{\sigma}_{t|T} r_{t-1}^{\gamma} \Big|, \tag{2.42}$$

$$\text{TIC} = \frac{\sqrt{1/T\sum_{t=1}^{T}\left(|y_t| - \hat{\sigma}_{t|T} r_{t-1}^{\gamma}\right)^2}}{\sqrt{1/T\sum_{t=1}^{T}|y_t|^2} + \sqrt{1/T\sum_{t=1}^{T}(\hat{\sigma}_{t|T} r_{t-1}^{\gamma})^2}}, \tag{2.43}$$

where $\gamma = 0$ for the crude oil model. For all tests, a lower value indicates a better model fit although RMSE and MAE depends on the scale used whereas TIC does not. An issue with these tests is that volatility is unobserved. By using the absolute residuals as a proxy for volatility, the tests may not tell us how accurately volatility is captured as the absolute residuals are at best a noisy estimate of volatility[12]. In this case we consider an additional test to evaluate the different SV model specifications by considering the empirical distribution of the standardized residuals[13]. The test can be described as follows - let us define $\hat{\epsilon}_t = y_t/\hat{\sigma}_{t|T} r_{t-1}^{\gamma}$ where $\hat{\sigma}_{t|T} = \mathbb{E}[\sigma_t|y_{1:T}]$ are the smoothed estimates outlined in Section 2.3. If the estimates of $\sigma_t$ are accurate under each procedure, then $\hat{\epsilon}$ should approximately follow a N(0,1) distribution. A standard normal distribution for the standardized residuals is determined using the Jarque-Bera test for normality and an examination of the QQ-plots.

The results of the likelihood ratio tests can be found in Table 2.5 whilst the remaining test results can be found in Table 2.6 for the interest rate data and Table 2.7 for the crude oil data.

**Short-Term Interest Rates**

As with the parameter estimates of the three estimation procedures, the model selection criteria show some differences. The LR tests found that none of the alternative unrestricted models should be chosen under QML estimation as opposed

[12]The tests were also conducted using the squared residuals as a proxy for volatility. As similar results were found to the absolute residuals, for brevity, the results have not been reported.

[13]A similar test was conducted in Liesenfeld and Richard (2003) for evaluating SV models in equity markets.

**Table 2.5**
Likelihood Ratio tests

This table provides LR tests for each model vs. the SV-AR model for each estimation procedure. The table reports the LR test statistics which are $\chi^2_{\text{DOF}}$ distributed where DOF denotes the degrees of freedom. P-values for the LR tests are in parentheses.

| | | Estimation Procedure | | |
|---|---|---|---|---|
| Model | DOF | QML | MCL | PF |
| Panel A: Short-Term Interest Rates | | | | |
| SV-ARMA | 1 | 0.3503 | 1.4999 | 0.7740 |
| | | (0.5539) | (0.2207) | (0.3790) |
| SV-ARX | 1 | 3.6022 | 10.7123 | 9.8574 |
| | | (0.0577) | (0.0011) | (0.0017) |
| SV-ARMAX | 2 | 3.8646 | 12.6655 | 11.1297 |
| | | (0.1448) | (0.0018) | (0.0038) |
| Panel B: Crude Oil | | | | |
| SV-ARMA | 1 | 1.7928 | 5.1030 | 3.9892 |
| | | (0.1806) | (0.0239) | (0.0458) |

to the standard AR(1) model. However, under both MCL and PF estimation, the LR tests indicate that both the ARX and ARMAX models are more favourable although the AR model can not be rejected in favour of the ARMA model.

The AIC and BIC results are consistent with the LR tests. Under QML, the AIC ranks the models (in descending order) as follows: ARX, AR, ARMAX, ARMA; under MCL, the AIC ranks are: ARX, ARMAX, AR, ARMA; and under PF the AIC ranks are: ARX, ARMAX, AR, ARMA. Using the BIC values, the ranks under QML are: AR, ARX, ARMAX, ARMA; under MCL: ARX, ARMAX, AR, ARMA; and under PF: ARX, AR, ARMAX, ARMA. The ARX model is chosen as the most appropriate model for all cases except for the BIC value under QML which chose the AR model. However, although there was agreement on the best performing model, there were differences between the rankings of the other models which again shows

**Table 2.6**

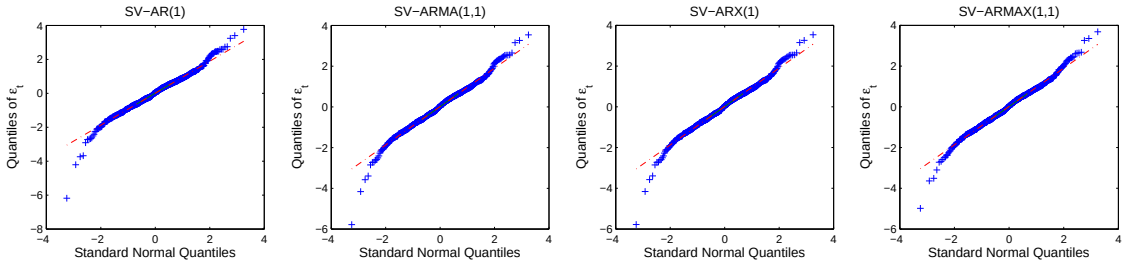
Model Diagnostics for Short-Term Interest Rates

This table provides diagnostic tests for the short-term interest rate data for each model and estimation procedure. AIC and BIC are the Akaike and Bayesian (Schwartz) information criterion. RMSE, MAE and TIC are computed according to (2.41), (2.42) and (2.43) respectively. JB is the Jarque-Bera test-statistic applied to the standardized residuals with p-values in parentheses.

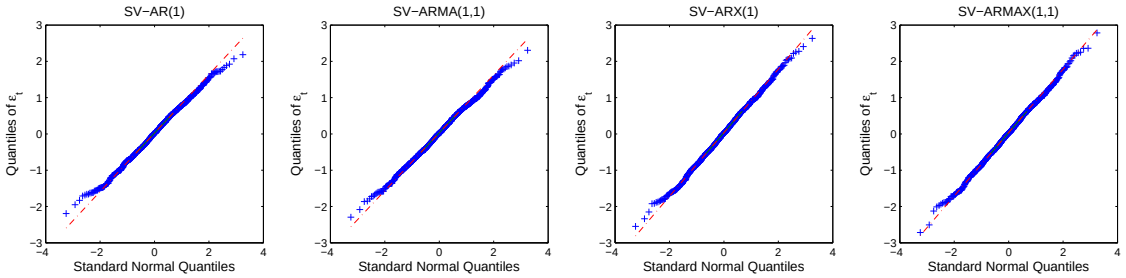
| Model | Loglik | AIC | BIC | RMSE | MAE | TIC | JB |
|---|---|---|---|---|---|---|---|
| QML | | | | | | | |
| SV-AR | −1994.58 | 3997.17 | 4016.17 | 0.0624 | 0.0429 | 0.0197 | 144.48 (<0.001) |
| SV-ARMA | −1994.40 | 3998.82 | 4022.56 | 0.0617 | 0.0423 | 0.0192 | 105.15 (0.000) |
| SV-ARX | −1992.78 | 3995.56 | 4019.31 | 0.0622 | 0.0428 | 0.0196 | 68.17 (<0.001) |
| SV-ARMAX | −1992.65 | 3997.30 | 4025.80 | 0.0608 | 0.0420 | 0.0186 | 49.93 (<0.001) |
| MCL | | | | | | | |
| SV-AR | 869.75 | −1731.50 | −1712.50 | 0.0702 | 0.0544 | 0.0212 | 4.67 (0.0891) |
| SV-ARMA | 870.44 | −1730.87 | −1707.12 | 0.0725 | 0.0559 | 0.0225 | 2.28 (0.3041) |
| SV-ARX | 875.11 | −1740.22 | −1716.47 | 0.0669 | 0.0506 | 0.0205 | 1.81 (0.3892) |
| SV-ARMAX | 876.08 | −1740.17 | −1711.67 | 0.0689 | 0.0515 | 0.0215 | 0.41 (>0.5) |
| PF | | | | | | | |
| SV-AR | 871.83 | −1735.66 | −1716.66 | 0.0575 | 0.0444 | 0.0154 | 12.69 (0.0049) |
| SV-ARMA | 872.22 | −1734.44 | −1710.68 | 0.0601 | 0.0454 | 0.0168 | 9.72 (0.0118) |
| SV-ARX | 876.76 | −1743.52 | −1719.77 | 0.0576 | 0.0447 | 0.0155 | 13.09 (0.0044) |
| SV-ARMAX | 877.40 | −1742.79 | −1714.29 | 0.0599 | 0.0456 | 0.0168 | 9.94 (0.0110) |

that the choice of estimation procedure may affect results.

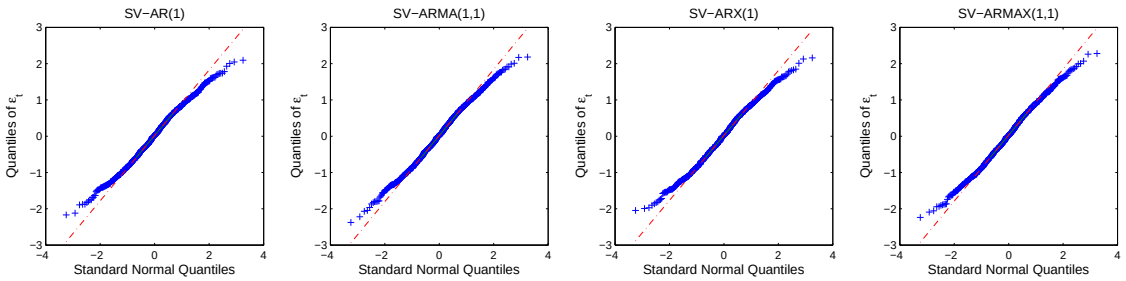
Comparing the model specifications for each estimation procedure, the RMSE, MAE and TIC results are fairly consistent with the AIC, BIC and LR test results. When comparing these metrics between the estimation procedures, it is interesting to note that PF is the best performing estimation procedure whereas MCL is the worst performing. However, both the QML and PF procedures indicate worse fits for the more specified models compared to SV-AR(1). On the other hand, under MCL the metrics indicate improving performance for the more specified models with SV-ARX(1) providing the best fit.




**(a)** QML




**(b)** MCL




**(c)** PF

**Figure 2.3**
QQ-plots of standardized residuals for short-term interest rates

If we consider the distributional properties, the JB test shows that under both QML and PF, the smoothed volatility estimates are not able to capture volatility very accurately as normality can be rejected for the standardized residuals of all models. On the other hand, under MCL estimation, normality can not be rejected for any of the models. The QQ-plots of the standardized residuals are shown in Figure 2.3. It can be seen that there is a significant degradation in the tails for the QML estimates, whereas the standardized residuals are much closer to a standard normal distribution for both the MCL and PF estimates. The plots suggest that the QML standardized residuals are much more dispersed than a standard normal distribution and that the estimated volatilities under QML are too smooth[14]. The opposite case is found for both MCL and PF. The QQ-plots for both MCL and PF show that the standardized residuals are less dispersed than standard normal random variables. This suggests that the models are over fitted under both MCL and PF estimation, although the degree of over fitting is much larger for PF. Examining the QQ-plots for MCL indicate that going from the SV-AR model to the SV-ARMAX model, adding both the MA and FOMC terms improve the volatility estimates as the standardized residuals progressively become closer to standard normal. This trend is not observed for the QQ-plots of the QML and PF procedures.

Interestingly, the RMSE, MAE and TIC values are not consistent with the what can be seen in the QQ-plots. In particular, although the QQ-plots demonstrate better performance of the MCL estimates, it was worst when considering the other metrics. Hence, this sample shows that although RMSE, MAE and TIC can give an indication of the relative size of the errors in the volatility estimates for the models, it does not give any information on whether volatility is accurately estimated.

[14]This is consistent with previous findings of the QML procedure, e.g. see Broto and Ruiz (2004).

**Crude Oil**

Looking at the crude oil results, they show similar characteristics to the interest rate results. The LR test shows that under both MCL and PF, the AR model can be rejected in favour of the ARMA model at the 5% level. However, under QML the AR model can not be rejected. For the AIC values, both MCL and PF ranks the ARMA model over AR, but QML chooses the AR model as more appropriate. For the BIC values, the higher penalty given to the extra parameter in the ARMA model results in the AR chosen over ARMA for all estimation procedures.

**Table 2.7**
Model Diagnostics for Crude Oil

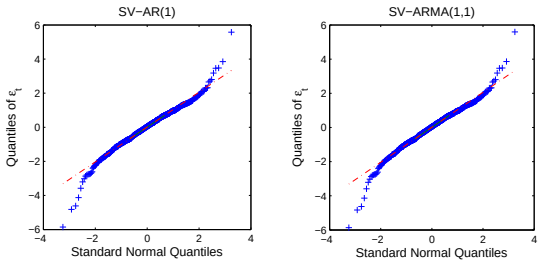
This table provides diagnostic tests for the crude oil data for each model and estimation procedure. AIC and BIC are the Akaike and Bayesian (Schwartz) information criterion. RMSE, MAE and TIC are computed according to (2.41), (2.42) and (2.43) respectively. JB is the Jarque-Bera test-statistic applied to the standardized residuals with p-values in parentheses.

| Model | Loglik | AIC | BIC | RMSE | MAE | TIC | JB |
|---|---|---|---|---|---|---|---|
| QML | | | | | | | |
| SV-AR | −1940.57 | 3887.15 | 3901.40 | 3.6896 | 2.5620 | 0.3650 | 247.80 (0.0000) |
| SV-ARMA | −1939.68 | 3887.36 | 3906.36 | 3.4648 | 2.4017 | 0.3420 | 129.66 (0.0000) |
| MCL | | | | | | | |
| SV-AR | −2553.43 | 5112.85 | 5127.11 | 3.5909 | 2.6023 | 0.3373 | 6.00 (0.0481) |
| SV-ARMA | −2550.88 | 5109.76 | 5128.76 | 3.3411 | 2.4055 | 0.3180 | 7.86 (0.0226) |
| PF | | | | | | | |
| SV-AR | −2554.50 | 5115.00 | 5129.25 | 3.4222 | 2.4600 | 0.3261 | 6.02 (0.0477) |
| SV-ARMA | −2552.50 | 5113.01 | 5132.01 | 3.2135 | 2.3067 | 0.3086 | 6.64 (0.0366) |

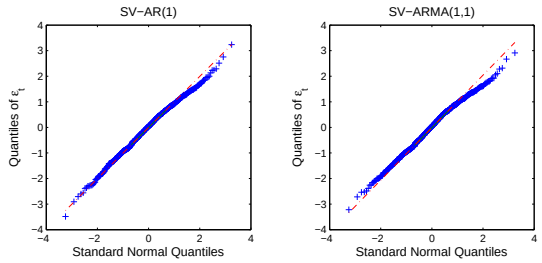
The RMSE, MAE and TIC values all indicate that the ARMA model improves model fit under all of the estimation procedures which somewhat agrees with the

results above. The final tests which investigate the distributional properties of the standardized residuals for each model shows that QML results in poor volatility estimates. The MCL and PF standardized residuals also reject normality at the 5% level of significance although significance is marginal and none of the models reject normality at the 1% level. Interestingly, the ARMA model appears to result in standardized residuals that are more non-normal for both MCL and PF. To get a better idea of the behaviour of the volatility estimated under each of the models and estimation procedures we look at the QQ-plots. As with the interest rate results, there is clear evidence that the QML estimates poorly capture crude oil volatility as the plot significantly deviates from standard normal. Both MCL and PF are much closer to the standard normal line for both the AR and ARMA models than QML. However closer inspection indicates more of a deviation from the standard normal line for the ARMA models similar to the interest rate results which suggest overfitting of volatility for the more specified ARMA model. Interestingly, the best fit according to the QQ-plots appears to be the MCL estimates of the AR model. The slightly poorer fit of the PF estimates appears to be as a result of trying to fit outliers. This is especially true of the PF estimates of the ARMA model where the interior of the plot seems to fit poorly but the two outliers seem to lie on the standard normal line.

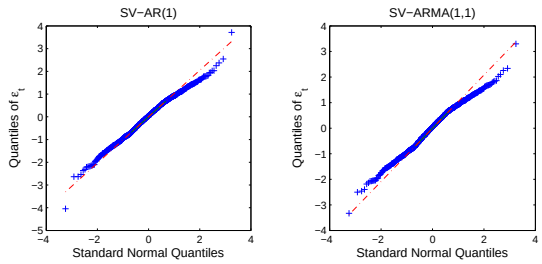
Combining the QQ-plots with the results of the JB tests tells us that the MCL procedure is able to more accurately estimate volatility than the two other procedures and that the SV model is appropriate for interest rate modelling when an efficient estimation procedure is used. These results demonstrate that choosing stochastic volatility models based on metrics which compare the absolute residuals or squared residuals with the volatility estimates of the different models should not be solely relied on. Although RMSE, MAE and TIC was able to agree with the other measures when comparing between models, there appear to be inconsisten-




**(a)** QML




**(b)** MCL




**(c)** PF

**Figure 2.4**
QQ-plots of standardized residuals for crude oil

cies when comparing between estimation procedures. Thus, it is recommended that additional diagnostic tests such as the one applied here which considers the distributional properties of the return series should be used to gauge the performance of SV models.

### 2.5.2 Forecasting Results

An important consideration of both the estimation procedure and the model choice is the out-of-sample forecasting performance. In this study, we investigate whether the estimation procedure makes any difference in forecasting performance. To maintain

a sufficient sample size for both estimation purposes and evaluating the out-of-sample forecasts, we choose a rolling window method where we fix the sample size for estimation to 654 observations and produce one-week ahead forecasts for 200 observations. The procedure is as follows, we estimate the model using the first 654 observations and forecast the 655th observation, then estimate the model using the next 654 observations starting from the 2nd observation and forecast the 656th observation, and so on until 200 forecasts are obtained.

To evaluate the forecasts, we again consider metrics similar to the in-sample tests. We compare the absolute residuals with the conditional volatility forecasts $\mathbb{E}[\sigma_{T+1|T}]$ by computing the RMSE, MAE and TIC as above. However, as mentioned previously, using the absolute residuals as a proxy for volatility does not give the full indication of model performance[15]. Hence we again conduct an additional test to determine the distributional properties of the residuals. Whilst we could attempt to construct a standardized residuals series similar to the in-sample test, this will not be appropriate in this case. This is because using the conditional mean of volatility to standardize the residuals will not result in standard normal random variables. This can be shown as follows,

$$\frac{y_{t+1}}{\mathbb{E}[\sigma_{t+1}|y_{1:t-1}]r_t^{\gamma}} = \frac{\epsilon_{t+1}\sigma_{t+1}}{\mathbb{E}[\sigma_{t+1}|y_{1:t-1}]},$$

where the right hand side is still a non-normal process.

An alternative is to consider the forecast density (FD) test of Berkowitz (2001) which uses the result from Rosenblatt (1952). This test is similar to the test considered by Kim et al. (1998) and Liesenfeld and Richard (2003) for stochastic volatility models

[15]As with the in-sample tests, the squared residuals were also used as a proxy for volatility, however the results remained the same as using absoluate residuals. They are thus not reported here.

but is not specific to the estimation procedures considered by these authors[16]. The FD test was introduced due to the inadequacy of value at risk (VaR) as a risk measure. This is because VaR can only detect whether forecasts lie within a certain forecast interval, whereas this test evaluates forecasting performance by considering the entire density of the forecasts. Consider the following transformation,

$$u_t = \int_{-\infty}^{y_t} \hat{f}_y(u|y_{1:t-1})du = \hat{F}_y(y_t|y_{1:t-1}), \tag{2.44}$$

where $y_t$ are the residuals defined by (2.3), $t = T+1, \ldots, T+R$, $\hat{f}_y(\cdot|y_{1:t-1})$ is the conditional probability density function (pdf) of $y_t$ given observations up to time $t-1$ and $\hat{F}_y(\cdot|y_{1:t-1})$ is the cumulative distribution function (cdf) of $y_t$. If the SV model is properly specified, then Rosenblatt (1952) showed that $u_t$ should be approximately uniformly distributed between 0 and 1. To evaluate the expression $u_t$, we need to derive the conditional cdf of $y_t$, $\hat{F}_y(\cdot|y_{1:t-1})$ which can be obtained using (2.3). If we define $e_t = y_t/r_{t-1}^{\gamma} = \sigma_t\epsilon_t$, then

$$\hat{F}_y(y_t|y_{1:t-1}) = F_e(e_t|y_{1:t-1}).$$

We show in Appendix 2.C that $F_e(\cdot|y_{1:t-1})$, is given by

$$F_e(e_t|y_{1:t-1}) = \begin{cases} \displaystyle\int_{-\infty}^{e_t}\int_{-\infty}^{0} \phi(u;0,1)\phi\left(\log\left(\frac{v}{u}\right)^2;\bar{x}_{t|t-1},\sigma_\eta\right)\frac{1}{|u|}dudv, \text{ if } e_t < 0, \\ \displaystyle\frac{1}{2} + \int_{0}^{e_t}\int_{0}^{\infty} \phi(u;0,1)\phi\left(\log\left(\frac{v}{u}\right)^2;\bar{x}_{t|t-1},\sigma_\eta\right)\frac{1}{u}dudv, \text{ if } e_t \geq 0, \end{cases} \tag{2.45}$$

where the expression can be evaluated using numerical integration routines available in Matlab. Having obtained $u_t$, evaluating the forecast density requires testing the

[16]Kim et al. (1998) requires application of a particle filter and Liesenfeld and Richard (2003) requires output of their importance sampler to apply the tests. The test here still relies on the output of the filters but does not rely on the estimation procedure.

sequence for uniformity. However, it is much easier to test for normality than uniformity and Berkowitz (2001) shows that if the sequence, $u_t$, is transformed by applying a standard normal inverse transform, the transformed sequence retains the same inaccuracies of the density forecast as the original sequence. This means we can apply standard tests of normality when testing the transformed sequence. Thus we transform $u_t$ as follows: $z_t = \Phi^{-1}(u_t)$ where $\Phi^{-1}(\cdot)$ is the standard normal inverse function. The model is correctly specified if $z_t$ is a sequence of serially uncorrelated standard normal random variables. To determine the distributional properties of $z_t$, Berkowitz (2001) suggests the following LR test. Consider the following AR(1) model for $z_t$,

$$z_t - \mu = \rho z_{t-1} + \nu_t, \tag{2.46}$$

where $\nu_t \sim$ i.i.d. $N(0, \sigma)$ distributed. If $z_t$ are serially uncorrelated standard normal random variables, it implies that $\mu = 0$, $\sigma = 1$ and $\rho = 0$. Hence the LR test involves testing the null hypothesis that $\mu = 0$, $\sigma = 1$ and $\rho = 0$, against the alternative of no restrictions to $\mu$, $\sigma$ and $\rho$. Berkowitz (2001) provides the likelihood function for the AR(1) model and the unrestricted parameter estimates can be found by maximizing this likelihood function. If we denote the log-likelihood function for the AR(1) model by $L(\mu, \sigma, \rho)$, then the log-likelihood function for the AR(1) model given by (2.5.2) is

$$\begin{aligned} L(\mu, \sigma, \rho) = -\frac{1}{2}\Big\{ & T\log(2\pi) + \log(\sigma^2/(1-\rho^2)) + \frac{[z_{T+1} - \mu/(1-\rho^2)]^2}{2\sigma^2(1-\rho^2)} \\ & + (T-1)\log(\sigma^2) + \sum_{t=T+2}^{T+R} \frac{(z_t - \mu - \rho z_{t-1})^2}{2\sigma^2} \Big\}. \end{aligned} \tag{2.47}$$

The LR test statistic is then given by

$$\text{LR} = -2(L(0, 1, 0) - L(\hat{\mu}, \hat{\sigma}, \hat{\rho})), \tag{2.48}$$

where $\hat{\mu}$, $\hat{\sigma}$ and $\hat{\rho}$ are the maximum likelihood estimates of $\mu$, $\sigma$ and $\rho$ respectively, and LR is a $\chi^2(3)$ distributed random variable. However, this test only has power to detect non-normality in the first two moments, and so in addition to this LR test we also determine normality of $z_t$ by applying the Jarque-Bera (JB) test. The JB test detects non-normality from the sample skewness and kurtosis. By jointly applying the LR and JB test, we are able to determine whether the models are misspecified.

**Short-Term Interest Rates**

The results of the out-of-sample forecast tests for the interest rate data can be found in Panel A of Table 2.8. If we look at the RMSE, MAE and TIC values, then there appears to be little distinguishing the estimates under the three procedures although QML appears to have the lowest RMSE, MAE and TIC values whereas PF has the largest. This result disagrees with the in-sample results where PF had the lowest values for these metrics giving more of an indication that this approach is not appropriate. Hence to gain a better idea of how well the procedures perform out-of-sample we consider the FD test results.

Looking at the distributional properties of the forecasts, we first consider the JB test of the transformed variables. Although not reported here, the results of the JB tests found that normality could not be rejected for any of the estimation procedures and model specifications. This allows us to proceed with the forecast density LR tests. Similar to the in-sample results, QML performs quite poorly as an i.i.d. standard normal distribution is strongly rejected for all model specifications. However, both MCL and PF appear to perform quite well as an i.i.d. standard normal distribution can not be rejected at the 5% level for all models except for the SV-ARX model under MCL. The forecasts generated under PF also appear to perform slightly

**Table 2.8**
Out-of-sample tests

This table reports the out-of-sample forecast tests for each model and estimation procedure. Panel A contains the tests for short-term interest rate data and Panel B contains the tests for the crude oil data. Each panel is divided into sub-tables by model type. RMSE, MAE and TIC are computed according to (2.41), (2.42) and (2.43) respectively. JB is the Jarque-Bera test-statistic. FD-LR is the forecast density likelihood ratio test statistic computed using (2.48). The p-values for the FD-LR test are in parentheses.

Panel A: Short-term interest rates

| SV-AR(1) | | | | SV-ARMA(1,1) | | | |
|---|---|---|---|---|---|---|---|
| | Estimation Procedure | | | | Estimation Procedure | | |
| | QML | MCL | PF | | QML | MCL | PF |
| RMSE | 0.0429 | 0.0460 | 0.0471 | RMSE | 0.0428 | 0.0484 | 0.0474 |
| MAE | 0.0332 | 0.0369 | 0.0380 | MAE | 0.0331 | 0.0402 | 0.0384 |
| TIC | 0.0160 | 0.0169 | 0.0174 | TIC | 0.0160 | 0.0180 | 0.0176 |
| JB | 0.35 | 0.07 | 0.28 | JB | 0.39 | 0.68 | 0.06 |
| | (>0.5) | (>0.5) | (>0.5) | | (>0.5) | (>0.5) | (>0.5) |
| FD-LR | 120.61 | 7.36 | 6.19 | FD-LR | 124.99 | 7.05 | 6.21 |
| | (0.0000) | (0.0613) | (0.1028) | | (0.0000) | (0.0701) | (0.1018) |

| SV-ARX(1) | | | | SV-ARMAX(1,1) | | | |
|---|---|---|---|---|---|---|---|
| | Estimation Procedure | | | | Estimation Procedure | | |
| | QML | MCL | PF | | QML | MCL | PF |
| RMSE | 0.0440 | 0.0476 | 0.0486 | RMSE | 0.0439 | 0.0496 | 0.0490 |
| MAE | 0.0341 | 0.0377 | 0.0386 | MAE | 0.0340 | 0.0401 | 0.0388 |
| TIC | 0.0167 | 0.0179 | 0.0185 | TIC | 0.0166 | 0.0189 | 0.0186 |
| JB | 0.38 | 0.88 | 0.11 | JB | 0.43 | 1.77 | 0.32 |
| | (>0.5) | (>0.5) | (>0.5) | | (>0.5) | (0.3632) | (>0.5) |
| FD-LR | 130.77 | 7.85 | 6.44 | FD-LR | 134.60 | 7.09 | 6.41 |
| | (0.0000) | (0.0492) | (0.0922) | | (0.0000) | (0.0689) | (0.0933) |

Panel B: Crude Oil

| SV-AR(1) | | | | SV-ARMA(1,1) | | | |
|---|---|---|---|---|---|---|---|
| | Estimation Procedure | | | | Estimation Procedure | | |
| | QML | MCL | PF | | QML | MCL | PF |
| RMSE | 3.2263 | 3.2722 | 3.3916 | RMSE | 3.2263 | 3.2923 | 3.4730 |
| MAE | 2.5843 | 2.6535 | 2.7723 | MAE | 2.5803 | 2.6668 | 2.8539 |
| TIC | 0.3815 | 0.3605 | 0.3335 | TIC | 0.3584 | 0.3351 | 0.3080 |
| JB | 13.44 | 7.59 | 6.29 | JB | 8.79 | 7.55 | 5.79 |
| | (0.0073) | (0.0274) | (0.0404) | | (0.0200) | (0.0277) | (0.0480) |
| FD-LR | 268.80 | 4.06 | 4.16 | FD-LR | 225.64 | 4.44 | 4.66 |
| | (0.0000) | (0.2553) | (0.2444) | | (0.0000) | (0.2182) | (0.1985) |

better than MCL as the LR test values are slightly lower. In terms of comparing the model specifications, a formal comparison between models can not be conducted using the FD test statistic. However, lower values of the FD statistic indicate that the transformed variables are closer to an N(0,1) distribution. In this sense, the FD tests indicate that forecasting performance degrades with the more specified models under QML and PF. However for MCL, it appears that the SV-ARX(1) model performs the worst out-of-sample and the SV-ARMA(1,1) performs the best although again there is little distinguishing the models.

**Crude Oil**

The results of the crude oil forecast tests can be found in Panel B of Table 2.8. The results show some similar characteristics with the out-of-sample interest rate results. For instance, the QML forecasts performed better than both MCL and PF according to the RMSE, MAE and TIC values. These are also inconsistent with the in-sample results where PF performed the best according to these metrics. However, both the JB test and the forecast density LR test both reject standard normality of $z_t$ for both the AR and ARMA models. This indicate that the forecasts estimated under QML poorly capture future volatility. For MCL and PF, the JB test is rejected at the 5% level for all of the models, although similar to the in-sample results normality can not be rejected at the 1% level. An inspection of $z_t$ and the returns $r_t$ found the presence of two large negative outliers. Exluding either of these observations results in non-rejection of normality for all of the tests which suggests that using more observations may be needed to determine how well the forecasts perform. As the FD-LR tests found that it can not be rejected that $z_t$ is serially uncorrelated with a mean of 0 and standard deviation of 1, it suggests a longer forecasting window may reveal whether volatility is captured well by the

SV models. The FD-LR tests therefore indicate that volatility forecasted with the parameters estimated under MCL and PF performed better than QML. However, this result is not conclusive since normality is rejected for all estimation procedures.

Overall, both MCL and PF have fairly similar results, with the density forecast test results suggesting that the SV model captures the volatility of short-term interest rates and crude oil quite well out-of-sample. Combining the in-sample and out-of-sample results, it appears that MCL produces the most efficient estimates of volatility and whilst PF performed slightly better out-of-sample, the in-sample results were worse than MCL.

## 2.6 Conclusion

This chapter has conducted a comparison of three maximum likelihood estimation procedures, QML, MCL and PF, for the estimation of stochastic volatility models using interest rate and crude oil data. Whilst it is clear that all models can be estimated freely using the three estimation procedures, there are advantages and disadvantages between them. The consistency of the QML estimator is questionable for finite samples although it is by far computationally the most efficient and the estimator is asymptotically normally distributed. From the empirical results it appears that MCL estimation provides the most robust results as the procedure searches for a Gaussian importance density close to the true posterior density. However, whilst the popularity of the particle filter has increased substantially recently, the same can not be said about the MCL procedure. One possible reason for this is the difficulty in determining an appropriate importance density under MCL estimation. The choice of importance density is crucial in its performance and can

be difficult (if it exists) to derive. It also must be obtained through an iterative procedure when attempting to match the modes of the importance density to the posterior density which increases the difficulty of implementation. On the other hand, although choosing an optimal importance density can improve performance in PF methods, the choice is not as crucial as MCL. A popular choice of importance density is the prior distribution which leads to straightforward calculations when updating the weights and propagating the filter forward.

A finite sample study of the three procedures found that QML exhibited slightly more bias than either MCL or PF. In the empirical application, the QML parameter estimates differed from the MCL and PF estimates which were quite close. Both the in- and out-of-sample goodness-of-fit tests demonstrate some mixed results, which shows a number of different tests should be considered when choosing the best model. In terms of in-sample performance, it was found that the QML procedure is not accurate enough when estimating volatility compared to MCL and PF. When comparing the performance of extensions to the SV model, although the tests indicated some improvement from the more specified models, the differences were not substantial. For interest rates, the results are consistent with the view that short-term interest rates follow a CIR model with stochastic volatility. They also showed that meetings of the FOMC significantly affect interest rate volatility, which suggests that the actions of central banks should be considered when modelling interest rate volatility. For crude oil, the standard AR(1) specification was found to be sufficient as it is not significantly outperformed by the ARMA(1,1) specification.

In respect of the out-of-sample results, there appears to be little distinguishing the different SV model specifications. However, the MCL and PF estimation results suggest that volatility is captured reasonably well by this class of stochastic volatility models. The main suggestions out of this study is that when choosing between

model specifications for univariate SV models, a number of different criteria should be considered and that the estimation procedure is important in accurately recovering the latent volatility process. In evaluating the performance of the models and estimation procedures, it was found that using the absolute residuals as a proxy for volatility in goodness-of-fit tests is not appropriate when evaluating volatility models, an approach taken by a few previous studies. The results showed that the RMSE, MAE and TIC values indicate QML has comparable or better performance than both MCL and PF in both in- and out-of-sample tests despite demonstrating much worse performance when comparing the distributional properties. This indicates that evaluating volatility models using absolute or squared residuals as a proxy for volatility should be supplemented by additional tests such as the forecast density test used in this study. In terms of the choice of a maximum likelihood estimation procedure, the MCL procedure performed the best both in-sample and out-of-sample. However, in applications where an importance density does not exist for the MCL procedure, the PF is a viable alternative.

# Appendices

## Appendix 2.A MCL Algorithm

As described in Section 2.3.2, the approximating Gaussian model is chosen to coincide at the mode of the true density. We determine the approximating model using the same approach as Koopman and Lee (2004) for the SV model given by Equation (2.3) and (2.4). The conditional log-density function for the true model is

$$p_t = \log p(y_t|x_t) = -\frac{1}{2}\left(\log(2\pi) + 2\gamma\log(r_{t-1}) + x_t + \frac{y_t^2}{r_{t-1}^{2\gamma}e^{x_t}}\right), \tag{2.49}$$

and for the approximating model

$$q_t = \log q(y_t|x_t) = -\frac{1}{2}\left(\log(2\pi b_t^2) + \frac{(y_t - a_t - x_t)^2}{b_t^2}\right). \tag{2.50}$$

The first and second derivatives with respect to $x_t$ for the two density functions are given by

$$\begin{aligned} p_t' &= \frac{1}{2}(y_t^2/r_{t-1}^{2\gamma}e^{x_t} - 1) & q_t' &= (y_t - a_t - x_t)/b_t^2, \\ p_t'' &= -\frac{1}{2}y_t^2/r_{t-1}^{2\gamma}e^{x_t} & q_t'' &= -1/b_t^2, \end{aligned}$$

and equating gives

$$b_t^2 = 2r_{t-1}^{2\gamma}e^{x_t}/y_t^2 \qquad\qquad y_t - a_t = x_t + 1 - \frac{1}{2}b_t^2. \tag{2.51}$$

As the solutions of both $a_t$ and $b_t^2$ require the latent variable $x_t$ which is not known initially, the solutions need to be approximated. Koopman and Lee (2004) suggest

this is achieved in practice by choosing a trial $x$ and estimating the values of $a_t$ and $b_t^2$ using the equations above. Given the approximating model, a new $x$ can be extracted using the Kalman filter and smoother. These steps are repeated until either $a_t$, $b_t^2$ or $x$ converge. Convergence is generally quite fast with around 10 iterations required for the changes in either $a_t$ or $b_t^2$ to be less than $1e^{-7}$. Once the approximating model has been found, the simulation smoother of Durbin and Koopman (2002) can be applied to make draws from $q(x|y)$. Readers are referred to that paper for details of implementing the simulation smoother.

The algorithm for the MCL procedure is as follows,

1. Choose a trial $x$
2. Use equation (2.51) to generate $b_t^2$ and $y_t - x_t$
3. Apply the Kalman filter and smoother to the approximating model with $\tilde{y}_t = y_t - a_t$
4. Obtain the smoothed state $\hat{x}_t = \mathbb{E}_{q(\cdot|y)}[x_t|y]$ for $t = 1, \ldots, T$. Set $x = \hat{x}$ and repeat steps 2 & 3 until convergence.
5. Simulate from the importance density $q(x|y)$ using a simulation smoother.

# Appendix 2.B PF Algorithm

When applying the smooth particle filter for parameter estimation, it is assumed that the importance density is the prior, $q(x_t|y_t, x_{t-1}) = p(x_t|x_{t-1})$. The algorithm for the smooth particle filter of Pitt (2002) is given as follows,

1. Initialize at $t = 0$. For $i = 1, \ldots, N$, sample $x_0^{(i)}$ from the prior $p(x_0)$.

2. For $t = 1, \ldots, T$ and $i = 1, \ldots, N$

   - Sample $x_t^{(i)}$ from $q(x_t|x_{0:t-1}^{(i)}, y_{1:t}) = p(x_t|x_{t-1}^{(i)})$, $i = 1, \ldots, N$.
   - Sort particles.
   - Evaluate the importance weights

$$w_t^{(i)} = w_{t-1}^{(i)} p(y_t|x_t^{(i)}).$$

   - Normalize importance weights

$$\tilde{w}_t^{(i)} = \frac{w_t^{(i)}}{\sum_{j=1}^{N} w_t^{(j)}}.$$

   - Construct smooth empirical cdf of sorted particles using linear interpolation. Resample from smooth cdf. Set weights of resampled particles to 1/N.

# Appendix 2.C Conditional PDF of the SV model

First we need to determine the pdf of $\sigma_t|y_{1:t-1} = \exp(x_t/2)|y_{1:t-1}$, which we use the following result. For a random variable $x$ with pdf $f_x(x)$, the pdf of $y = g(x)$, $f_y(y)$ is given by

$$f_y(y) = \frac{f_x(g^{-1}(y))}{g'(g^{-1}(y))}. \tag{2.52}$$

From Equation (2.4), we have $f_{x_t}(x_t|y_{1:t-1}) = \phi(x_t; \bar{x}_{t|t-1}, \sigma_\eta)$ where $\phi(\cdot; \mu, \sigma)$ denotes the normal pdf with mean $\mu$ and variance $\sigma^2$ and $\bar{x}_{t|t-1} = \mathbb{E}[x_t|y_{1:t-1}]$. By setting

$g(x) = \exp(x/2)$, we have $g'(x) = \exp(x/2)/2$ and $g^{-1}(y) = \log y^2$. Letting $x = x_t|y_{1:t-1}$ and $y = \sigma_t|y_{1:t-1}$, the conditional pdf of $\sigma_t$ is given by,

$$f_\sigma(\sigma_t|y_{1:t-1}) = \frac{2\phi(\log\sigma_t^2; \bar{x}_{t|t-1}, \sigma_\eta)}{\sigma_t}. \tag{2.53}$$

To derive the conditional pdf, $e_t|y_{1:t-1}$, we need to determine the pdf of the product of the two random variables $\epsilon_t|y_{1:t-1}$ and $\sigma_t|y_{1:t-1}$. This is derived by using the result of Rohatgi (1976) who shows that for two random variables x and y, with joint density $f_{x,y}(x, y)$, the pdf of $v = xy$ is

$$f_v(v) = \int_{-\infty}^{\infty} f_{x,y}\Big(u, \frac{v}{u}\Big)\frac{1}{|u|}du. \tag{2.54}$$

We note that under the assumption that $\epsilon_t$ and $\eta_t$ are independent implies that $\epsilon_t$ and $\sigma_t$ are also independent and so the joint density of $\epsilon_t$ and $\sigma_t$ is

$$f_{\epsilon,\sigma}(\epsilon_t, \sigma_t|y_{1:t-1}) = f_\epsilon(\epsilon_t|y_{1:t-1})f_\sigma(\sigma_t|y_{1:t-1}). \tag{2.55}$$

This means the conditional pdf of $e_t$ is given by,

$$\begin{aligned} f_e(e_t|y_{1:t-1}) &= \int_{-\infty}^{\infty} f_\epsilon(u|y_{1:t-1})f_\sigma\Big(\frac{e_t}{u}|y_{1:t-1}\Big)\frac{1}{|u|}du \\ &= \begin{cases} \displaystyle\int_{-\infty}^{0} \phi(u;0,1)\phi\left(\log\Big(\frac{e_t}{u}\Big)^2; \bar{x}_{t|t-1}, \sigma_\eta\right)\frac{1}{|u|}du, & \text{if } e_t < 0 \\ \displaystyle\int_{0}^{\infty} \phi(u;0,1)\phi\left(\log\Big(\frac{e_t}{u}\Big)^2; \bar{x}_{t|t-1}, \sigma_\eta\right)\frac{1}{u}du, & \text{if } e_t \geq 0 \end{cases}, \end{aligned} \tag{2.56}$$

where $\phi(\cdot;\mu,\sigma)$ is the pdf of a normally distributed random variable with mean $\mu$ and standard deviation $\sigma$. Given the pdf above, we obtain the expression for $F_e(e_t|y_{1:t-1})$ given by Equation (2.45) by integrating the density, $f_e(\cdot|y_{1:t-1})$ over the

domain $[-\infty, e_t]$.

# Chapter 3

# Time-varying Market Price of Risk in the Crude Oil Futures Market[17]

## 3.1 Introduction

Trading in commodity futures markets has grown considerably over the last decade with crude oil being the most actively traded commodity in the world. It has become increasingly important to accurately model crude oil prices and a number of different modelling approaches have been proposed in the literature. However, due to empirical evidence suggesting that commodities exhibit mean-reversion (e.g. Schwartz (1997)), the modelling approaches have differed somewhat from that of stock markets. One of the main approaches of modelling commodity prices advocates the theory of storage where spot prices are driven by convenience yields and interest rates such as Schwartz (1997), Routledge et al. (2000) and Casassus and Collin-Dufresne (2005). However, as convenience yields are not directly identifiable or tangible streams of income, the concept proves difficult to interpret. An alternative approach considers the empirical properties of commodities where spot prices may deviate from some equilibrium level in the short term but tend to mean-revert

[17]A shortened version of this chapter is forthcoming in the *Journal of Futures Markets*. The article is currently available for early view on the journal's website. The article reference is: Bhar, R., Lee, D., 2010. Time-varying market price of risk in the crude oil futures market. *Journal of Futures Markets*, Forthcoming. doi: 10.1002/fut.20493.

to this equilibrium level, which is itself uncertain. This approach was introduced by Schwartz and Smith (2000) (SS hereafter) and has been a popular approach due to its intuitive appeal for the modelling of commodity spot prices.

In this chapter and the next, the latter approach to commodity modelling is taken where a three-factor short/long term factor model is considered. The approach is similar to Cortazar and Naranjo (2006) and Dempster et al. (2008) and extends the SS short/long term two-factor model for the modelling of commodity prices. Under this modelling approach commodity prices are decomposed into two main components. One component represents the long-term equilibrium price level, whereas the other component represents short-term deviations from the equilibrium price. The long-term component is modelled by a geometric Brownian which represents the uncertainty in the equilibrium price level due to factors such as technological advancements, major supply discoveries or exhaustion, and regulatory or political factors. The short-term component is modelled by two Ornstein-Uhlenback processes which represent short-term fluctuations in prices. These fluctuations could be due to changes in supply and demand caused by short-term supply disruptions, or fluctuations in the cost of production such as storage, transportation and financing costs. A side effect of the mean-reversion in prices is that the models are consistent with the "Samuelson effect" (Samuelson (1965)) where the volatility of futures prices close to maturity are higher than the volatilities of futures contracts with longer maturities.

As the model is an affine-diffusion model, it allows for tractable solutions to derivative prices. In this study, we price futures contracts under the three-factor model. To derive the futures price, we take a no-arbitrage approach where futures prices are considered to be the expectation of future spot prices under a "risk neutral" or "equivalent martingale measure". In terms of specifying a risk-neutral mea-

sure, the main consideration is the market price of risk (MPR) specification. The methodology involves postulating a model for the spot price under the real-world measure and by no-arbitrage arguments assuming a form of the MPR to obtain a model under the risk-neutral measure. Whilst the most popular specification for the MPR is a constant in most applications, there is evidence to suggest that incorporating a time-varying MPR can better explain the distribution of asset prices. Cheridito et al. (2007) estimated a number of one-, two- and three-factor affine term structure models with different specifications of the MPR and found that incorporating time-varying MPR improved the time-series fit of the models. Casassus and Collin-Dufresne (2005) estimated a three-factor stochastic convenience yield model for commodity futures prices that incorporates a time-varying MPR specification in the convenience yield factor that depends on the spot price and interest rates. They found some evidence that time-varying risk premiums exist in crude oil, copper, gold and silver futures markets.

Other authors consider variations on the SS short/long term model to estimate the market price of risk. Kolos and Ronn (2008) estimated the sign and magnitude of the market price of risk for energy futures prices under the assumption that the MPR was constant. They found that the MPR for most energy commodities were positive, reflecting the state of normal backwardation exhibited by futures prices. Weron (2008) investigated the MPR implied by Asian-style options and futures in the Nordic electricity market and found that it varied significantly over time. Cartea and Williams (2008) investigated time-varying MPR under the short/long term model with seasonality in the natural gas market although no conclusive finding could be made due to the high standard error of the time-varying MPR parameter.

In terms of the crude oil futures market, allowing risk premiums to be time-varying is well motivated by the literature, albeit from a different modelling perspective. For

instance, Deaves and Krinsky (1992), Moosa and Al-Loughani (1994) and Considine and Larson (2001) found evidence for time-varying risk premiums by analyzing the returns on futures contracts. Both Deaves and Krinsky (1992) and Moosa and Al-Loughani (1994) tested whether the efficient market hypothesis holds in the crude oil futures market. Moosa and Al-Loughani (1994) found that market efficiency should be rejected whilst Deaves and Krinsky (1992) could not find conclusive evidence that the market is not efficient, although both suggest the results are due to time-varying risk premiums. Considine and Larson (2001) finds evidence for time-varying risk premiums in the crude oil and natural gas markets and found that risk premiums rose sharply with greater price volatility.

Thus to allow risk premiums to be time-varying under the three-factor model, we allow the MPR specification to be time-varying. In defining the risk-neutral measure, we follow a similar framework to Duffee (2002) and Cheridito et al. (2007) and assume an MPR that is a linear function of the state variables. This allows the MPR to be time-varying and also has the implication that the mean-reversion of the two short-term factors is different under the real-world and risk-neutral measures. The model is estimated using data consisting of weekly observations of futures prices with contract maturities ranging from 1 month to 5 years allowing information from both cross-sectional and time-series data to be used in estimation. The Kalman filter is employed to estimate the model parameters by maximum likelihood and evaluate how well the three-factor model fits the term structure of futures prices compared to nested specifications. This chapter also compares the constant and time-varying MPR specifications and to determine whether there is evidence suggesting that time-varying risk premiums exist in the crude oil futures market. The forecasting performance is evaluated and whether a time-varying MPR specification outperforms a constant MPR specification. However, as the choice of MPR specification is based only on the statistical fit of the model, further investigation

of the risk premiums is needed.

Given the estimated parameters of the model, an estimate of the latent time-varying market price of risk is extracted, which enables us to determine whether any relationship exists between the risk premiums in the crude oil futures market and risk factors in the macroeconomy. A number of studies have linked time-varying risk premiums in commodity futures markets to the macroeconomy although not much work has involved energy markets. Bessembinder and Chan (1992) found that macroeconomic variables such as Treasury bill yields, equity dividend yields and 'junk' bond premium possess forecast power in agricultural, metals and currency futures markets. Bailey and Chan (1993) found that the variability in the basis of a number of agricultural and metal commodities can be attributed to time-varying risk premiums and that this variability is correlated with risk factors common to stock and bond markets such as credit spreads and stock dividend yields. Baum and Barkoulas (1996) conducted a similar study for the currency futures basis where they found time-varying risk premiums exist and can be forecast by dividend yields, default spreads and term spreads. More recently, Sadorsky (2002) found that four macroeconomic variables have significant forecast power in explaining time-varying risk premiums in petroleum futures markets and improve the out-of-sample forecasting performance. These studies estimate risk premiums using either ex-post returns as the difference between futures prices and realized future spot prices or excess returns from holding futures contracts. They also only consider contracts with one or two maturities and hence do not incorporate the entire term structure of futures prices when estimating the relationships. For this thesis, we follow a similar approach to these studies by conducting a regression analysis of the MPR and the risk factors described above. The findings suggest that although crude oil futures prices are affected by idiosyncratic risk factors, risk premiums in crude oil markets are driven by the same factors as equity and bond markets.

The contributions of this chapter are as follows. Firstly, by considering a three-factor model for commodity prices, the model is able to model the term structure of futures prices more accurately than previous studies that consider two-factor short/long term models. The model also incorporates a time-varying MPR specification that is found to better explain the time-series behaviour of crude oil prices than under a constant MPR specification. As the data set includes futures prices with maturities ranging from 1 month to 5 years this ensures that the risk premiums are representative of a range of futures contracts with varying maturities. Furthermore, whilst most authors make certain assumptions about the MPR specification when pricing derivatives, there is little economic justification for this choice apart from the fact that the models admit no arbitrage. For instance, although some of the modelling approaches in the literature mentioned above incorporate a time-varying MPR, the risk premiums themselves are not investigated in much detail. As many MPR specifications satisfy the no-arbitrage assumption, it is not clear which MPR specification is appropriate. Additionally, given that the model is able to accurately capture futures prices, we turn to the MPR assumption and explore whether the risk premiums have any economic justification. As such, this chapter investigates whether the financial mathematical approach of modelling crude oil spot prices and the assumptions made about risk premiums when pricing derivatives is consistent with what is observed in the macroeconomy. Furthermore, whilst a number of studies have investigated systematic risk factors in commodity markets, few have included energy markets and thus this study adds to the literature in this regard by investigating the crude oil market.

This chapter is structured as follows. Section 3.2 introduces the model and the time-varying market price of risk specification. Section 3.3 discusses the role of risk premiums in the crude oil market and provides some evidence of their existence. Section 3.4 provides a description of the data and estimation procedure. Section 3.5

presents the empirical results. Section 3.6 discusses the time-varying risk premiums estimated in this model. Section 3.7 concludes the chapter.

## 3.2 Short/Long Term Factor Model

The short/long factor model was introduced by SS and has proven quite popular as a basis for commodity pricing models in the literature due to its simple and intuitively appealing interpretation. The model separates the spot price into two components - a short-term factor which captures any short-term deviations in prices as a result of short-term supply and demand fluctuations; and a long-term factor which represents the equilibrium price level and captures permanent changes in prices related to the cost of production. To name a few papers using the approach, Sørensen (2002) and Cartea and Williams (2008) consider the SS model with the inclusion of seasonality for pricing commodity futures contracts and Aiube et al. (2008) extend the model by including a jump process in the short-term component.

However, although the model is intuitively appealing, the original SS model consists of only two factors resulting in poor performance in capturing the term structure of futures prices for a wide range of maturities. To improve the modelling performance, some authors have extended the two-factor model to include additional factors. For instance, Dempster et al. (2008) proposed a three-factor short/long factor jump-diffusion model and Cortazar and Naranjo (2006) generalized the SS model to an N-factor model. In both cases, it was found that at least two mean-reverting factors are required to accurately model the term structure of futures prices whilst Cortazar and Naranjo (2006) found marginal improvement in futures pricing fit when including a fourth factor although the model is better able to capture the volatility

term structure. This study takes a similar approach by considering a three-factor model with two mean-reverting short-term factors and one long-term factor but incorporates a time-varying MPR specification.

In our model specification, we assume that the spot price $S_t$ is defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ where the filtration $\mathcal{F} = \{\mathcal{F}_t\}_{t\geq 0}$ is generated by a 3-dimensional Brownian motion $Z_t$. The spot price is described by a three-factor model and is given by the following set of equations,

$$\begin{aligned} \log S_t &= X_{1,t} + X_{2,t} + X_{3,t}, \\ dX_{1,t} &= -\kappa_1 X_{1,t} dt + \sigma_1 dZ_{1,t}, \\ dX_{2,t} &= -\kappa_2 X_{2,t} dt + \sigma_2 dZ_{2,t}, \\ dX_{3,t} &= \mu_3 dt + \sigma_3 dZ_{3,t}, \end{aligned} \tag{3.1}$$

where $Z_{i,t}, i = 1, 2, 3$ are correlated Brownian motions under $\mathbb{P}$ and $dZ_{i,t}dZ_{j,t} = \rho_{ij}dt$, for $i, j = 1, 2, 3$, $i \neq j$. The model allows for non-stationary long-term prices due to $X_{3,t}$ consistent with the recent trend in commodity markets where prices have been increasing. Also, as commodities are consumption assets, $X_{1,t}$ and $X_{2,t}$ allow prices to exhibit a supply-demand equilibrium in the short-term.

This approach of modelling commodity prices is in contrast to convenience yield models which attempt to capture the most important features consistent with the theory of storage as advocated by Kaldor (1939), Working (1949), Brennan (1958), Deaton and Laroque (1996) and Routledge et al. (2000). These authors found that convenience yields arise endogenously as a result of the interaction between supply, demand and storage decisions. The three-factor convenience yield models of Schwartz (1997) and Casassus and Collin-Dufresne (2005) model the relationship between these factors by a single stochastic process which represents the convenience

yield net of storage costs. Some authors also introduce an exogenous stochastic interest rate process to model commodity prices. However, as noted by Dempster et al. (2008), in empirical studies where stochastic interest rates are introduced, the volatility of the interest rate process is very low compared to the volatility of the convenience yield and spot price process. This is the case found in Schwartz (1997) and Casassus and Collin-Dufresne (2005) where the volatility of the interest rate factor is found to be less than 5% the size of the volatility of the convenience yield and spot price process with little improvement in model fit[18].

Although the model does not explicitly take account of convenience yield and storage costs in its dynamics, the two-factor model has been shown by SS to be equivalent to the Gibson and Schwartz (1990) model that incorporates stochastic convenience yields. Furthermore, Dempster et al. (2008) demonstrate that a three-factor short/long factor model is equivalent to a two-factor convenience yield model where the cost of carry is modelled by two factors. In effect, the model captures the interaction between supply, demand and storage decisions using two factors rather than one factor. The model is therefore able to remain consistent with the "theory of storage" whilst retaining an intuitively appealing interpretation and accurately modelling futures prices[19]. A further appeal of this modelling approach is that the mean-reverting short-term factor allows futures price to exhibit the "Samuelson Effect" (Samuelson (1965)). This is where the volatility of futures prices decreases with increasing time to maturity and is a property found in a number of commodity markets.

[18]Both Schwartz (1997) and Casassus and Collin-Dufresne (2005) estimated the volatilities of the interest rate process to be around 0.008-0.009 in comparison to the volatilities of the spot price and convenience yields of over 0.3. In addition, Schwartz (1997) did not find much evidence of an improvement in model fit over the model without the interest rate factor.

[19]Whilst we could consider the equivalent convenience yield model instead, we remain with this three-factor model as it allows for a simpler time-varying market price of risk specification. Also, convenience yields are not tangible income streams and representing the relationship between supply, demand and storage decisions directly seems more intuitively appealing.

To price futures and options contracts the usual approach in the literature is to use the risk-neutral pricing framework. For diffusion models, pricing under the risk-neutral measure usually involves adjusting the drift of the diffusion process by a constant amount where the drift adjustment is termed the market price of risk. Recently, some papers have considered incorporating a time-varying MPR approach such as Casassus and Collin-Dufresne (2005) and Cartea and Williams (2008) in commodity markets or Duffee (2002) and Cheridito et al. (2007) in the fixed income market where the MPR is specified as linear functions of the state variables. This approach is applied to the short/long factor model allowing for a more flexible MPR specification which can capture difference in the time-series properties of the model under the real world and risk-neutral measures.

Assuming that the market is arbitrage-free and complete, by the fundamental theorem of asset pricing of Harrison and Pliska (1981) an equivalent martingale measure (EMM) or risk-neutral measure, $\mathbb{Q}$, exists. For the model studied in this chapter, whilst the MPR is allowed to be time-varying, it is assumed that the model remains structurally of the same class under both $\mathbb{P}$ and $\mathbb{Q}$[20]. Under this assumption, only the two short-term factors are allowed to have time-varying MPR specifications whereas the long-term factor MPR remains a constant. To derive the risk-neutral measure, we consider a different representation for the model involving standard Brownian motions. If $Z_t = (Z_{1,t}, Z_{2,t}, Z_{3,t})'$ are correlated Brownian motions, then (3.1) can be re-expressed in terms of standard Brownian motion as follows,

$$dX_t = (\mu + \kappa X_t)dt + \Sigma dW_t, \tag{3.2}$$

[20]This means that both $X_{1,t}$ and $X_{2,t}$ remain OU processes under $\mathbb{P}$ and $\mathbb{Q}$ whereas $X_{3,t}$ remains a Brownian motion under $\mathbb{P}$ and $\mathbb{Q}$.

where

$$\mu = \begin{pmatrix} 0 \\ 0 \\ \mu_3 \end{pmatrix}, \ \kappa = \begin{pmatrix} -\kappa_1 & 0 & 0 \\ 0 & -\kappa_2 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

$$\Sigma = \begin{pmatrix} \Sigma_{11} & 0 & 0 \\ \Sigma_{21} & \Sigma_{22} & 0 \\ \Sigma_{31} & \Sigma_{32} & \Sigma_{33} \end{pmatrix} = \begin{pmatrix} \sigma_1 & 0 & 0 \\ \rho_{12}\sigma_2 & \sqrt{1-\rho_{12}^2}\sigma_2 & 0 \\ \rho_{13}\sigma_3 & \frac{\rho_{23}-\rho_{13}\rho_{12}}{\sqrt{1-\rho_{23}^2}}\sigma_3 & \sqrt{1-\rho_{12}^2-\frac{\rho_{23}-\rho_{13}\rho_{12}}{\sqrt{1-\rho_{23}^2}}}\sigma_3 \end{pmatrix},$$

and $W_t = (W_{1,t}, W_{2,t}, W_{3,t})'$ is a 3-dimensional standard Brownian motion under $\mathbb{P}$.

Assuming that the market is arbitrage-free and complete, by the fundamental theorem of asset pricing of Harrison and Pliska (1981) an equivalent martingale measure (EMM) or risk-neutral measure, $\mathbb{Q}$, exists. We define the change of measure by the Radon-Nikodym derivative,

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = \exp\Big(-\int_0^T \gamma_t \cdot dW_t - \frac{1}{2}|\gamma_t|^2 dt\Big) = \eta_T, \tag{3.3}$$

where $|\cdot|$ is the Euclidean norm and $\gamma_t$ is the MPR defined as

$$\gamma_t = \Sigma^{-1}(\alpha + \beta X_t), \tag{3.4}$$

where

$$\alpha = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} \quad \text{and} \quad \beta = \begin{pmatrix} \beta_1 & 0 & 0 \\ 0 & \beta_2 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

The MPR is therefore a linear combination of each state variable. By assuming

sufficient regularity so that Novikov's condition holds in some finite positive time interval, i.e.

$$\mathbb{E}\left[\exp\Big(\frac{1}{2}\int_0^T |\gamma_t|^2 dt\Big)\right] < \infty,$$

then $\eta_T$ is a martingale and $\mathbb{E}(\eta_T) = 1$. By Girsanov's theorem, $dW_t^* = dW_t + \gamma_t dt$ is a 3-dimensional standard Brownian motion under $\mathbb{Q}$. The dynamics of $X_t$ under the risk-neutral measure become

$$dX_t = (\mu^* + \kappa^* X_t)dt + \Sigma dW_t^*, \tag{3.5}$$

where $\mu^* = \mu - \alpha$ and $\kappa^* = \kappa - \beta$. Making the transformation back to correlated Brownian motion $Z_t$, the risk neutral dynamics of the model is

$$\begin{aligned} dX_{1,t} &= \big(-\alpha_1 - \kappa_1^* X_{1,t}\big)dt + \sigma_1 dZ_{1,t}^*, \\ dX_{2,t} &= \big(-\alpha_2 - \kappa_2^* X_{2,t}\big)dt + \sigma_2 dZ_{2,t}^*, \\ dX_{3,t} &= \mu_3^* dt + \sigma_3 dZ_{3,t}^*, \end{aligned} \tag{3.6}$$

where $Z_{i,t}, i = 1, 2, 3$ are correlated Brownian motions under $\mathbb{Q}$ with $dZ_{i,t}dZ_{j,t} = \rho_{ij}dt$, for $i, j = 1, 2, 3$, $i \neq j$, $\mu_3^* = \mu_3 - \alpha_3$, $\kappa_1^* = \kappa_1 + \beta_1$ and $\kappa_2^* = \kappa_2 + \beta_2$. It should be noted that under the short/long term model, the risk-free rate is assumed to be constant and implicit in the risk premium. It is therefore not explicitly specified.

Given that the short-term factors are a linear combination of each factor, if we set $\beta_1$ and $\beta_2$ to zero, then the model reverts to a constant MPR specification. Under the terminology introduced by Duffee (2002), the time-varying MPR model specification is a special case of the essentially affine class whereas the constant MPR model specification belongs to the completely affine class[21]. Furthermore, as

[21] According to Duffee (2002), a requirement for the completely affine class is that variation in the market price of risk depends entirely on volatility. As volatility is constant in this model,

volatility is assumed constant in the model, if we denote the instantaneous risk premium of the model by $\Lambda_t$, then

$$\begin{aligned}\Lambda_t &= \lambda_{1,t} + \lambda_{2,t} + \lambda_{3,t} \\ &= \alpha_1 + \beta_1 X_{1,t} + \alpha_2 + \beta_2 X_{2,t} + \alpha_3.\end{aligned} \tag{3.7}$$

The risk premium is therefore also a linear combination of the two short-term factors.

To derive the futures price, it is noted that the solutions to the set of stochastic differential equations (SDE) for $X_{1,t}$, $X_{2,t}$ and $X_{3,t}$ can be easily found as they are jointly normally distributed. As the log of the spot price is just the sum of the factors, it is also normally distributed. This means the conditional mean and variance of $\ln S_T$ under the risk neutral measure is given by

$$\begin{aligned}\mathbb{E}_{\mathbb{Q}}[\log S_T|\mathcal{F}_t] = {} & e^{-\kappa_1^*(T-t)}X_{1,t} - \frac{\alpha_1}{\kappa_1^*}\big(1 - e^{-\kappa_1^*(T-t)}\big) + e^{-\kappa_2^*(T-t)}X_{2,t} \\ & - \frac{\alpha_2}{\kappa_2^*}\big(1 - e^{-\kappa_2^*(T-t)}\big) + X_{3,t} + (\mu_3 - \alpha_3)(T-t),\end{aligned} \tag{3.8a}$$

$$\begin{aligned}\mathrm{Var}_{\mathbb{Q}}[\log S_T|\mathcal{F}_t] = {} & \big(1 - e^{-2\kappa_1^*(T-t)}\big)\frac{\sigma_1^2}{2\kappa_1^*} + \big(1 - e^{-2\kappa_2^*(T-t)}\big)\frac{\sigma_2^2}{2\kappa_2^*} + \sigma_3^2(T-t) \\ & + \big(1 - e^{-(\kappa_1^*+\kappa_2^*)(T-t)}\big)\frac{2\rho_{12}\sigma_1\sigma_2}{\kappa_1^*+\kappa_2^*} + \big(1 - e^{-\kappa_1^*(T-t)}\big)\frac{2\rho_{13}\sigma_1\sigma_3}{\kappa_1^*} \\ & + \big(1 - e^{-\kappa_2^*(T-t)}\big)\frac{2\rho_{23}\sigma_2\sigma_3}{\kappa_2^*},\end{aligned} \tag{3.8b}$$

where $\mathbb{E}_{\mathbb{Q}}[\cdot]$ and $\mathrm{Var}_{\mathbb{Q}}[\cdot]$ are the mean and variance under the risk-neutral probability measure $\mathbb{Q}$.

Given that the futures price at time t is the expected value of the future spot price at time T under the risk-neutral measure (Cox et al. (1981)), the futures price,

then the MPR must be constant to be completely affine. Under the essentailly affine class, this restriction is relaxed and the market price of risk is allowed to be a linear combination of the state variables.

$F(t,T)$, is defined by

$$F(t,T) = \mathbb{E}_{\mathbb{Q}}[S_T|\mathcal{F}_t]. \tag{3.9}$$

As $\ln S_T$ is normally distributed, $S_T$ is log-normally distributed and the expected spot price under $\mathbb{Q}$ is given by

$$\mathbb{E}_{\mathbb{Q}}[S_T|\mathcal{F}_t] = \exp\Big\{ E_{\mathbb{Q}}[\log S_T|\mathcal{F}_t] + \frac{1}{2}\mathrm{Var}_{\mathbb{Q}}[\log S_T|\mathcal{F}_t]\Big\}. \tag{3.10}$$

Substituting (3.8a) into (3.10) and (3.9), the futures price, $F(t,T)$, is therefore given by

$$F(t,T) = \exp\Big(A^*(T-t) + B^*(T-t)'X_t\Big), \tag{3.11}$$

where

$$\begin{aligned} A(\tau) = (\mu_3 - \alpha_3)\tau - \frac{\alpha_1}{\kappa_1^*}\big(1-e^{-\kappa_1^*\tau}\big) - \frac{\alpha_2}{\kappa_2^*}\big(1-e^{-\kappa_2^*\tau}\big) + \frac{1}{2}\Bigg[(1-e^{-2\kappa_1^*\tau})\frac{\sigma_1^2}{2\kappa_1^*} \\ + \big(1-e^{-2\kappa_2^*\tau}\big)\frac{\sigma_2^2}{2\kappa_2^*} + \sigma_3^2\tau + \big(1-e^{-(\kappa_1^*+\kappa_2^*)\tau}\big)\frac{2\rho_{12}\sigma_1\sigma_2}{\kappa_1^*+\kappa_2^*} \\ + \big(1-e^{-\kappa_1^*\tau}\big)\frac{2\rho_{13}\sigma_1\sigma_3}{\kappa_1^*} + \big(1-e^{-\kappa_2^*\tau}\big)\frac{2\rho_{23}\sigma_2\sigma_3}{\kappa_2^*}\Bigg], \end{aligned}$$

and $B(\tau) = (e^{-\kappa_1^*\tau}, e^{-\kappa_2^*\tau}, 1)'$. A few models are nested under this specification. For instance, the constant MPR model of SS corresponds to having no $X_{2,t}$ factor and setting $\beta_1 = 0$ or the 2-factor time-varying MPR model considered by Cartea and Williams (2008) without seasonality is equivalent to the model without $X_{2,t}$. Whilst we are concerned with the risk premiums implied by the model, we also estimate the nested two-factor model as a comparison. In particular, we investigate whether the MPR specifications are plausible in either model.

## 3.3 Risk premiums in the crude oil market

As there are a number of different interpretations of risk premiums in the literature for commodities, we firstly define how risk premiums should be interpreted. The price risk premium, $\Pi_t$, is defined as the difference between the expected spot price under the real-world measure and the futures price,

$$\Pi_t = E_{\mathbb{P}}[S_T|\mathcal{F}_t] - F(t,T). \tag{3.12}$$

A similar definition is given in Geman (2005) and Weron (2008). Using this definition, the risk premium can be considered as the compensation required for holding the risky asset (commodity) over the risk-free asset (forward/futures contract). When risk premiums are positive, the market is said to be in backwardation and when risk premiums are negative the market is in contango. The theory of backwardation originated in Keynes (1930) where it was suggested that spot prices must exceed futures prices by the amount which producers are prepared to sacrifice to hedge themselves from the risk of a decrease in prices.

The evidence for risk premiums in commodity futures markets is mixed as noted by Fama and French (1987). However, there are a number of reasons for normal backwardation occurring in the crude oil futures market. As crude oil is a consumption commodity, long hedgers (i.e. companies or entities that hold crude oil in their inventory) do so for its consumption value and not for its investment value. Such entities would be unwilling to hedge the price of oil by selling oil and buying futures contracts. Also, as the short-term demand elasticity of crude oil is known to be very low and the ability for some to pass on any increase in the price of oil to their customers (e.g. airlines introduced a fuel levy due to the recent high

oil prices) reduces the need for hedging. Hence long hedgers may be able to afford short-term increases in prices and will only hedge using futures contracts if the price is favourable for them to do so. On the other hand, short hedgers do not have the same flexibility and it would be expected that hedgers in the market would be net short forward/futures contracts and it is speculators who make up the shortfall in buyers. Hence as speculators are essentially providing insurance for short hedgers, they require a risk premium to compensate them for their position resulting in futures prices lower than current spot prices.

Risk premiums could also arise due to a mismatch in the timing of hedging requirements for producers as opposed to consumers (including wholesalers and retailers). For producers, the capital required to fund the initial exploration and construction of production facilities requires long-term future cash flows to meet financing demands. Producers would therefore want to sell long dated forward/futures contracts to reduce the risk of adverse movements in oil prices. On the other hand, long hedgers such as wholesalers would not have as large capital requirements and would only want to enter shorter term forward/futures contracts for hedging purposes. Hence whilst hedgers exist on both sides of futures transactions for short dated contracts, it is generally only producers that require long dated contracts. As a result, it is speculators who buy the long dated contracts and again require a risk premium for providing this insurance to producers resulting in lower prices for longer dated futures contracts.

Evidence for the above can be found over the last 10 years, where spot oil prices have exhibited consistent growth and the market has remained in backwardation for much of the period. Whilst crude oil is a non-replenishable resource and has been viewed as an increasingly scarce resource for some time now, the 5 year futures contract usually does not trade above the nearby contract. This suggests that whilst

spot prices have exhibited a positive trend, the futures price for long-term contracts incorporate a risk premium as they have generally traded below the nearby contract. This can be clearly seen if we consider a plot of the futures prices for 1 month, 1 year and 5 year futures contracts in Figure 3.1 over the period February 1999 to August 2007. During the sub-period from 1999 - mid 2005, the market exhibited strong backwardation as the 5 year prices plotted below the 1 year prices and well below the 1 month prices for most of the period. Even from mid 2005 onwards, the 5 year futures price has remained quite close to both the 1 month and 1 year prices. In fact over the entire period, the average differential of the 5 year and 1 year contract prices over the 1 month contract price was -\$4.79 and -\$0.44 respectively. This is consistent with the view that risk premiums are paid by short hedgers to speculators to insure them against future price decreases and where the longer the maturity, the higher the risk premiums paid.

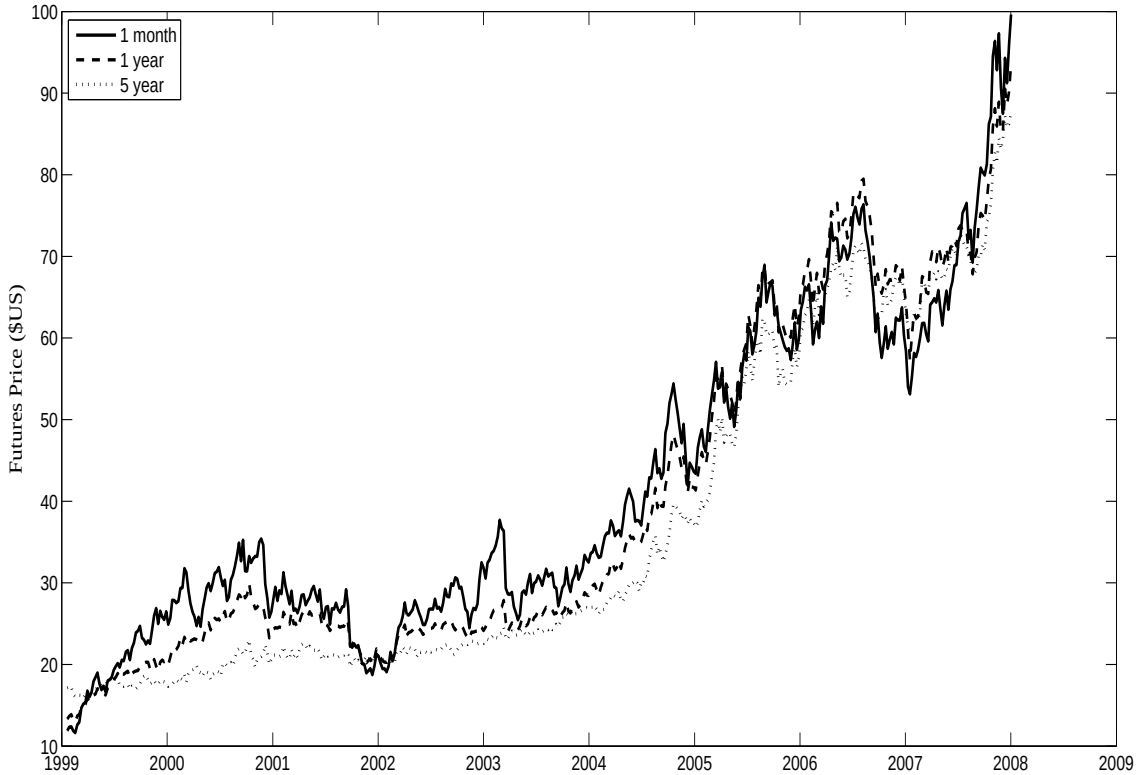



**Figure 3.1**
Plot of 1-month, 1 year and 5 year futures prices, Jan 1999 - Jan 2008

In our analysis of risk premiums in the crude oil market, we investigate the market

price of risk implied by the model introduced in Section 3.2. A few authors (e.g. Weron (2008), Kolos and Ronn (2008) and Cartea and Williams (2008)) have estimated different MPR specifications in energy markets although the risk premiums themselves are only investigated quite briefly. Weron (2008) estimated the MPR implied by Asian options in electricity markets using a mean-reverting jump diffusion model and described the evolution of MPR over time. Kolos and Ronn (2008) and Cartea and Williams (2008) estimated the MPR implied by the SS model. Kolos and Ronn (2008) assumed a constant MPR and estimated its sign and magnitude over different periods for crude oil and other energy commodities. Cartea and Williams (2008) estimated a time-varying MPR for the UK gas market but only discuss the sign and magnitude of the MPR over the period of study. They found no conclusive evidence about whether time-varying risk premiums exist in the UK gas market as the standard errors on the time-varying MPR parameters were not statistically significant. One of the possible reasons for the poor result is that the sample size and cross section of maturities they considered was quite small.

## 3.4 Data and Methodology

### 3.4.1 Data

The data consists of weekly observations of crude oil futures prices on the New York Mercantile exchange (NYMEX) obtained from Datastream. In order to obtain a substantial number of futures contracts with sufficient liquidity, the observation dates span from January 1999 to January 2008. Trading of these contracts ends on the third business day prior to the 25th calendar day preceding the delivery month. If the 25th is not a business day, trading ends on the fourth business day prior to

the last business day before the 25th calendar day. For the purpose of pricing the contracts, the maturity date of each contract is assumed to occur on the first day of the delivery month where we have chosen contracts of 9 different maturities with maturities of approximately 1 month, 3 months, 6 month, 1 year, 18 months, 2 years, 3 years, 4 years and 5 years.

As a note, every maturity month up to 60 months in the future were only available from the start of 2006. Prior to 2006, only maturities for the first 30 months are available, with subsequent contracts having maturities at approximately 6 month intervals up to 7 years in the future. For contracts with no exactly matching maturity, we choose the closest maturing contract. The total number of observations in the data set is $468 \times 9 = 4212$, which allows both the information from the time-series of crude oil futures prices and the cross-sectional data to be used for parameter estimation. The data is also split into a sub-sample from Janauary 1999 - January 2006 to use for diagnostic testing in-sample with a hold out period from January 2006 - January 2008 to conduct out-of-sample tests.

It should also be noted that for the nearby maturity contract, open interest decreases rapidly during the last couple of weeks prior to the last trading day[22]. To mitigate the effects of the drop in open interest, only futures contracts with at least 10 trading days prior to the last trading day are included in the analysis.

Descriptive statistics for the data is provided in Table 3.1. As can be observed from the mean of the futures price, the market has mainly been in backwardation as higher prices are observed for shorter maturity contracts as opposed to longer maturity contracts. Interestingly, the standard deviation of longer term futures prices are generally higher than the standard deviation of shorter term futures prices.

[22]The last trading day of the NYMEX crude oil futures contract occurs on the third business day prior to the twenty-fifth calendar day of the month preceding the delivery month.

**Table 3.1**
Descriptive Statistics for WTI Crude Oil Futures Contracts
This table reports the mean and standard deviation of the futures price and the annualized weekly returns (in percentages) of the West Texas Intermediate (WTI) crude oil futures contracts covering the period 6 January 1999 - 2 January 2008. The data set consists of 468 weekly observations of 9 futures contracts with the mean contract maturity (in months) reported in the second column.

| Contract | Mean Maturity (years) | Futures Price | | Futures Return (%) | |
|---|---|---|---|---|---|
| | | Mean | Standard Deviation | Mean | Standard Deviation |
| 1 | 0.0994 | 41.2087 | 19.4143 | 3.2849 | 31.5411 |
| 2 | 0.2380 | 41.1450 | 19.9062 | 3.2570 | 29.1394 |
| 3 | 0.4897 | 40.6401 | 20.5006 | 3.1706 | 24.6295 |
| 4 | 0.9934 | 39.5954 | 21.0877 | 3.0109 | 20.2823 |
| 5 | 1.4965 | 38.7567 | 21.2480 | 2.8795 | 18.2543 |
| 6 | 1.9999 | 38.1093 | 21.1781 | 2.7678 | 17.3604 |
| 7 | 2.9958 | 37.2465 | 20.9123 | 2.6435 | 17.4445 |
| 8 | 3.9872 | 36.7209 | 20.6377 | 2.5684 | 17.4747 |
| 9 | 4.9871 | 36.3504 | 20.4172 | 2.5201 | 17.8006 |

However, this does not reflect the volatility of futures prices on a weekly basis but rather the deviation of the price from the mean over the sample period. A better measure of futures price volatility is the standard deviation of the weekly returns where it is clear that the shorter maturity contract are more volatile than longer term contracts. It is also clearly seen in Figure 3.1 that the 1 month futures contract is much more volatile than 1 year and 5 year contracts. This demonstrates that the "Samuelson effect" is present in the crude oil futures market.

### 3.4.2 Estimation Procedure

There are two main approaches taken in the literature for calibrating the parameters of the model and estimating the MPR. One approach involves estimating the parameters of the model under the real-world measure using data from the underlying asset. Then whilst holding the parameters that are the same under both measures

constant, the MPR is estimated using either futures or options data by minimizing a loss function such as the sum of squared errors. This is the usual approach when estimating models for assets where both spot prices and futures or options are easily observed. For instance, Weron (2008) use this approach for estimating the MPR in electricity markets and Cartea and Williams (2008) in the natural gas market. However, for many commodities including crude oil, spot prices are not directly observable and most studies take the near month futures contract as a proxy for the spot price.

Another approach involves estimating the parameters under both the real world and risk neutral measures simultaneously when spot prices are not available. This is the approach taken by Schwartz (1997), Schwartz and Smith (2000) and Casassus and Collin-Dufresne (2005) to estimate the parameters in commodity models, Cheridito et al. (2007) to determine MPR specifications in a number of multi-factor term structure models and Kolos and Ronn (2008) to estimate the MPR in a number of energy markets using a term structure of forward and futures prices. Similarly, Pandher (2001) used estimating functions to estimate the MPR parameters from options on the S&P500 index.

As spot prices are not readily available in the crude oil market, the latter approach is taken. Parameter estimation is carried out using the Kalman filter in a similar manner to Schwartz (1997), Schwartz and Smith (2000) and Cortazar and Naranjo (2006). The Kalman filter is a recursive procedure for estimating the unobserved state variables from the observations given a particular state space model. In order to facilitate use of the Kalman filter, we need to cast the model into a state space form consisting of a state equation and a measurement equation. The state equation describes the evolution of the state variables to future states whilst the measurement equation describes the relationship between the observations and the latent states.

For the models considered here, the unobserved state variables are the three factors $(X_{1,t}, X_{2,t}, X_{3,t})$ and the observations are the futures prices. As the Kalman filter is a discrete filter and the models above are continuous-time models, the evolution of the state variables is described assuming discrete time steps corresponding to the exact solutions of the SDE given by (3.1).

It should be noted that the state equation depends on the transition density under the real-world measure $\mathbb{P}$. However, as we only observe futures prices that are priced under the risk-neutral measure $\mathbb{Q}$, the market price of risk parameters are not directly identifiable. To address this issue, a modified approach is taken to Schwartz (1997) and Schwartz and Smith (2000) where it is assumed that all futures prices are observed with measurement errors representing the errors from the reporting of prices or the estimation errors of the model. This allows the entire cross-section of futures prices to be used in estimating the time-series of the state variables with the Kalman filter. The transition density of the state variables can then be estimated from the time-series which also allows estimation of the model parameters under the real-world measure, $\mathbb{P}$. This includes estimating the market price of risk parameters.

However, rather than assuming that the errors are independent and identically distributed as in Schwartz (1997) and Schwartz and Smith (2000), a more robust approach involves assuming that the measurement errors follow independent Gaussian AR(1) processes as it allows us to detect for model misspecification. This is similar to the approach used in Casassus and Collin-Dufresne (2005) and Cheridito et al. (2007), although they do not use the Kalman filter for estimation. For simplicity, we assume that all errors have the same autoregressive factor, $\phi$, although they are allowed to have different variances as futures contracts of different maturities tend to exhibit idiosyncratic behaviour. The measurement errors therefore have the

following form,

$$\nu_{i,t} = \phi\nu_{i,t-1} + \epsilon_{i,t}, \qquad \epsilon_t \sim N(0, s_i^2), \tag{3.13}$$

for the $i^{\text{th}}$ futures contract where $i = 1, \ldots, 9$ denotes the number of each contract described in Table 3.1.

As the three-factor model and the nested two-factor model can be specified as linear Gaussian state space models, the likelihood functions can be calculated directly using the linear Kalman filter. This allows us to estimate the parameters by maximum likelihood estimation.

**State Space Model Estimation**

The state space model described above is now shown in more detail. Since the measurement errors follow an AR(1) process, they must be placed in the state equation. This results in the state space model having the following form,

$$\begin{aligned} y_t &= Z_t x_t + d_t, \\ x_t &= T_t x_{t-1} + c_t + \eta_t, \end{aligned} \tag{3.14}$$

where for $t = 1, \ldots, T$, $y_t$ are the $\mathbb{R}^n$ observation vectors, $x_t$ are the $\mathbb{R}^{(m+n)}$ state vectors and $\eta_t \sim N(0, Q_t)$ are the state errors. Given that there are $n$ futures contract maturities and $m = 3$ for the three-factor model, the variables in the state space model are defined as follows:
$y_t = \begin{pmatrix} \ln F(t, t+T_1) & \ldots & \ln F(t, t+T_n) \end{pmatrix}'$ is the vector of observed log futures prices with maturities $T_1, \ldots, T_n$,
$x_t = \begin{pmatrix} X_{1,t} & X_{2,t} & X_{3,t} & \nu_{1,t} & \ldots & \nu_{n,t} \end{pmatrix}'$ is the vector of the latent states representing

the two short-term and long-term factors and the measurement errors,

$Z_t = \begin{pmatrix} e^{-\kappa_1^* T_1} & e^{-\kappa_2^* T_1} & 1 & \mathbf{I}_n \\ \vdots & \vdots & \vdots & \\ e^{-\kappa_1^* T_n} & e^{-\kappa_2^* T_n} & 1 & \end{pmatrix}$ is the $\mathbb{R}^{n\times(m+n)}$ matrix that relates the latent states to the log futures price where $\mathbf{I}_n$ is the $n \times n$ identity matrix,

$d_t = \begin{pmatrix} A(T_1) & \dots & A(T_n) \end{pmatrix}'$, is a $\mathbb{R}^n$ vector defined in (3.9),

$T_t = \begin{pmatrix} e^{-\kappa_1 \Delta t} & 0 & 0 & \mathbf{0} \\ 0 & e^{-\kappa_2 \Delta t} & 0 & \mathbf{0} \\ 0 & 0 & 1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \phi \mathbf{I}_n \end{pmatrix}$ is the $\mathbb{R}^{(m+n)\times(m+n)}$ state transition matrix,

$c_t = \begin{pmatrix} 0 & 0 & \mu_3 \Delta t & 0 & \dots & 0 \end{pmatrix}'$ is a $\mathbb{R}^{3+n}$ vector,

$Q_t = \begin{pmatrix} \mathrm{Cov}\begin{pmatrix} X_{1,t} & X_{2,t} & X_{3,t} \end{pmatrix} & \mathbf{0} \\ \mathbf{0} & \mathrm{diag}\begin{pmatrix} s_1 & \dots & s_n \end{pmatrix} \end{pmatrix}$ is the covariance matrix of the state errors.

If the state space model matrices are examined, then it can be seen that the state transition equation depends only on the real-world parameters whereas the measurement equation depends only on the risk-neutral parameters. This allows both sets of parameters to be estimated from the state space model.

The two-factor models of Schwartz and Smith (2000), Sørensen (2002) and Cartea and Williams (2008) are nested in the three-factor model and can be obtained by restricting all parameters related to the second short-term factor, $X_{2,t}$ to zero. The parameter estimates are obtained by maximizing the (log-)likelihood function. The standard Kalman filter recursions introduced in Chapter 1, Section 1.2 can be applied in this case given that the model is linear and Gaussian.

In terms of parameter inference, standard errors are estimated by taking the in-

verse of the Hessian matrix which is computed using finite difference methods. All procedures were programmed in Matlab and the likelihood was maximized with the BFGS method used by the Matlab optimization routines. As there are a large number of parameters to be estimated, we used a number of different starting values to ensure that we have reached a global maximum rather than a local maximum.

## 3.5 Empirical Results

The maximum likelihood parameter estimates can be found in Table 3.2. The parameters estimated include the mean-reversion parameters under the risk-neutral measure for the two short-term factors, the drift of the long-term factor, the market price of risk parameters, the volatility and correlation parameters, and the volatility of the measurement errors. For the full three-factor model with time-varying MPR, a total of 23 parameters are estimated. In contrast, the nested constant and time-varying MPR models of SS have 16 and 17 parameters respectively to be estimated. The log-likelihood values are also included at the bottom of Table 3.2 allowing us to conduct likelihood ratio tests to compare the nested models. We now describe the parameter estimates for both the two-factor and three-factor models with and without a time-varying MPR.

### 3.5.1 Two-Factor Model

Considering the estimates under the two-factor model for the constant MPR specification, all parameters are highly significant except for $\alpha_1$ and $\rho_{13}$. If we consider the values of the parameters themselves, a few things should be noted. Firstly, $\sigma_1$

**Table 3.2**

Parameter Estimates

This table reports the parameter estimates for the two-factor and three-factor models obtained by maximum likelihood using the Kalman filter for the full-sample and a sub-sample period. Const denotes a constant MPR specification and TV denotes a time-varying MPR specification. The parameters are defined by equations (3.6), (3.7) and (3.13). Standard errors for the parameters are in parentheses.

| | Full Sample | | | | 1999-2006 | | | |
|---|---|---|---|---|---|---|---|---|
| | Two-Factor | | Three-Factor | | Two-Factor | | Three-Factor | |
| Parameter | Const | TV | Const | TV | Const | TV | Const | TV |
| $\kappa_1^*$ | 0.7349 | 0.7345 | 1.3300 | 1.3241 | 0.7471 | 0.7466 | 1.4555 | 1.4612 |
| | (0.0114) | (0.0114) | (0.0677) | (0.0679) | (0.0127) | (0.0127) | (0.0870) | (0.0877) |
| $\alpha_1$ | 0.2339 | 0.1916 | 0.0281 | 0.0188 | 0.2816 | 0.2286 | 0.0314 | 0.0126 |
| | (0.0930) | (0.0346) | (0.1018) | (0.1252) | (0.1111) | (0.0332) | (0.1129) | (0.1255) |
| $\beta_1$ | – | −1.3156 | – | 0.2437 | 0.0000 | −1.8641 | 0.0000 | 0.1642 |
| | – | (0.5381) | – | (0.4617) | (0.0127) | (0.6731) | (0.0870) | (0.6332) |
| $\sigma_1$ | 0.2780 | 0.2795 | 0.3006 | 0.3022 | 0.2928 | 0.2944 | 0.2979 | 0.2971 |
| | (0.0085) | (0.0086) | (0.0183) | (0.0192) | (0.0106) | (0.0107) | (0.0178) | (0.0181) |
| $\kappa_2^*$ | – | – | 0.6081 | 0.6134 | – | – | 0.6281 | 0.6318 |
| | – | – | (0.0265) | (0.0269) | – | – | (0.0277) | (0.0279) |
| $\alpha_2$ | – | – | 0.1806 | 0.1325 | – | – | 0.2419 | 0.1689 |
| | – | – | (0.0935) | (0.0340) | – | – | (0.1056) | (0.0421) |
| $\beta_2$ | – | – | – | −1.2227 | – | – | – | −1.1093 |
| | – | – | – | (0.4046) | – | – | – | (0.4657) |
| $\sigma_2$ | – | – | 0.2778 | 0.2823 | – | – | 0.2769 | 0.2789 |
| | – | – | (0.0171) | (0.0181) | – | – | (0.0160) | (0.0163) |
| $\mu_3$ | 0.1814 | 0.1814 | 0.1783 | 0.1781 | 0.1851 | 0.1850 | 0.1789 | 0.1785 |
| | (0.0586) | (0.0586) | (0.0597) | (0.0597) | (0.0681) | (0.0681) | (0.0693) | (0.0692) |
| $\mu_3^*$ | −0.0108 | −0.0108 | −0.0155 | −0.0155 | −0.0103 | −0.0103 | −0.0154 | −0.0153 |
| | (0.0017) | (0.0017) | (0.0016) | (0.0016) | (0.0019) | (0.0019) | (0.0019) | (0.0019) |
| $\sigma_3$ | 0.1755 | 0.1755 | 0.1789 | 0.1788 | 0.1798 | 0.1798 | 0.1832 | 0.1831 |
| | (0.0068) | (0.0068) | (0.0069) | (0.0069) | (0.0067) | (0.0067) | (0.0079) | (0.0079) |
| $\rho_{12}$ | – | – | −0.4144 | −0.4396 | – | – | −0.2995 | −0.3134 |
| | – | – | (0.0938) | (0.0954) | – | – | (0.0991) | (0.0994) |
| $\rho_{13}$ | −0.1636 | −0.1582 | 0.1614 | 0.1644 | −0.2156 | −0.2094 | 0.1483 | 0.1500 |
| | (0.0456) | (0.0457) | (0.0500) | (0.0500) | (0.0507) | (0.0508) | (0.0574) | (0.0574) |
| $\rho_{23}$ | – | – | −0.3097 | −0.3026 | – | – | −0.3545 | −0.3455 |
| | – | – | (0.0457) | (0.0458) | – | – | (0.0508) | (0.0511) |
| $\phi$ | 0.9712 | 0.9712 | 0.8729 | 0.8730 | 0.9702 | 0.9702 | 0.8783 | 0.8794 |
| | (0.0044) | (0.0044) | (0.0117) | (0.0117) | (0.0223) | (0.0223) | (0.0135) | (0.0135) |
| $s_1$ | 0.0164 | 0.0164 | 0.0106 | 0.0106 | 0.0174 | 0.0174 | 0.0114 | 0.0114 |
| | (0.0005) | (0.0005) | (0.0005) | (0.0005) | (0.0006) | (0.0006) | (0.0005) | (0.0005) |
| $s_2$ | 0.0067 | 0.0067 | 0.0034 | 0.0034 | 0.0072 | 0.0072 | 0.0035 | 0.0034 |
| | (0.0002) | (0.0002) | (0.0003) | (0.0003) | (0.0003) | (0.0003) | (0.0004) | (0.0004) |
| $s_3$ | 0.0000 | 0.0000 | 0.0018 | 0.0018 | 0.0000 | 0.0000 | 0.0020 | 0.0021 |
| | (0.0004) | (0.0004) | (0.0002) | (0.0002) | (0.0005) | (0.0005) | (0.0002) | (0.0002) |
| $s_4$ | 0.0041 | 0.0041 | 0.0015 | 0.0015 | 0.0043 | 0.0043 | 0.0016 | 0.0016 |
| | (0.0001) | (0.0001) | (0.0001) | (0.0001) | (0.0002) | (0.0002) | (0.0002) | (0.0002) |
| $s_5$ | 0.0051 | 0.0051 | 0.0022 | 0.0022 | 0.0052 | 0.0052 | 0.0025 | 0.0026 |
| | (0.0002) | (0.0002) | (0.0001) | (0.0011) | (0.0002) | (0.0002) | (0.0001) | (0.0001) |
| $s_6$ | 0.0057 | 0.0057 | 0.0041 | 0.0041 | 0.0060 | 0.0060 | 0.0046 | 0.0046 |
| | (0.0002) | (0.0002) | (0.0001) | (0.0001) | (0.0002) | (0.0002) | (0.0002) | (0.0002) |
| $s_7$ | 0.0029 | 0.0029 | 0.0023 | 0.0023 | 0.0030 | 0.0030 | 0.0025 | 0.0025 |
| | (0.0001) | (0.0001) | (0.0001) | (0.0001) | (0.0001) | (0.0001) | (0.0001) | (0.0001) |
| $s_8$ | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0005 | 0.0005 | 0.0000 | 0.0000 |
| | (0.0004) | (0.0004) | (0.0002) | (0.0002) | (0.0004) | (0.0004) | (0.0002) | (0.0002) |
| $s_9$ | 0.0026 | 0.0026 | 0.0025 | 0.0025 | 0.0028 | 0.0028 | 0.0027 | 0.0027 |
| | (0.0001) | (0.0001) | (0.0001) | (0.0001) | (0.0001) | (0.0001) | (0.0001) | (0.0001) |
| Loglik | 14739.38 | 14742.18 | 15611.53 | 15616.65 | 11308.65 | 11312.51 | 11913.15 | 11916.35 |

is much larger than $\sigma_3$ which indicates most of the variation in spot prices is due to the short-term factor. There is also some indication that significant risk premiums are present, especially for long-term contracts, as the value of $\mu_3$ is much larger than $\mu_3^*$ and $\alpha_1$ is quite large and significant. The values of $\mu_3$ and $\mu_3^*$ suggest that whilst long-term prices were expected to increase at a rate of around 18% each year under the real-world measure, this was not reflected in the futures price as the long-term risk-neutral drift was only 1%. Furthermore, the value of $\alpha_1 = 0.23$ suggests large risk premiums in short-term futures prices. The combined effect of the two results is that the futures market is generally in backwardation. This means that futures contracts usually trade at a lower price than the current spot price even though spot oil prices were generally rising over the period of study. This is consistent with what could clearly be seen in Figure 3.1. It is also consistent with the view that hedgers trading short positions in crude oil futures contracts pay a premium to speculators as they are essentially providing insurance to the hedgers for potential drops in oil prices.

In order to determine whether time-varying risk premiums exist in the crude oil futures markets we estimate the model incorporating time-varying MPR. For the two-factor model with a time-varying MPR in the short-term factor, it can be seen that the parameter estimates are very similar to the constant MPR model. All of the risk-neutral parameters differ only marginally from the estimates under the constant MPR model and they are all significant. This is similar to other papers that estimate time-varying MPR models who find that only the time-series properties of the underlying state variables are affected by incorporating a time-varying MPR[23].

If we consider the parameters related to the time-varying MPR, $\alpha_1$ is significant with

[23] As only changes in the risk-neutral parameters affect the cross-sectional fit of futures prices, this is not improved by incorporating time-varying MPR as mentioned by SS and is the reason for the minor improvement in the likelihood values.

a value of 0.19 which is slightly lower than the value estimated in the constant MPR model. The value of $\beta_1$ is $-1.32$ and is significant at the 5% level. The standard errors are relatively high compared to the other parameters which was also observed by SS and Cartea and Williams (2008) among others for MPR parameters. The main reason for this is the MPR parameters can not be directly identified from futures prices as the underlying spot price process is not directly observed. However, as this study has used a wider range of futures contracts and a longer time-series than these two studies, the level of statistical accuracy for $\beta_1$ is improved with evidence suggesting that a time-varying MPR is present. A likelihood ratio (LR) test is also applied to determine whether time-varying MPR exists. The results of the test can be found in the first set of columns under "Two-Factor" in Table 3.3. The parameter restriction in this case is $\beta_1 = 0$ and the LR test statistic is 5.6 which is higher than the $\chi^2_{0.05,1}$ critical value of 3.84. Thus under the two-factor model, there is evidence that a time-varying MPR exists in the crude oil market. The implications for the negative value of $\beta_1$ is that when the price of crude oil falls, risk premiums will be higher. This means that whilst the short-term price decrease may not be expected to last for long, this is not reflected in the futures price as speculators will tend to bid a lower price to hedgers.

However, although the two-factor model short/long model has been popular for modelling commodity prices, there is evidence that the model is misspecified here. The autocorrelation parameter of the measurement errors, $\phi$ is 0.97 which is very high. This suggests that the model is misspecified and that an additional factor may be required to correctly model futures prices.

**Table 3.3**
Likelihood Ratio tests

This table reports details of likelihood ratio (LR) tests conducted on the two and three-factor models for the full-sample and a sub-sample period. The columns labelled "Two-Factor" contain the LR tests of each model against the two-factor constant MPR model. The columns labelled "Three Factor" contain the LR test of the three-factor constant MPR against the three-factor time-varying MPR model. DOF denotes the degrees of freedom and LR denotes the LR test statistic which is a $\chi^2_{\text{DOF}}$ distributed. P-values for the LR tests are in parentheses.

| | Full-Sample | | | | 1999-2006 | | | |
|---|---|---|---|---|---|---|---|---|
| | Two-Factor | | Three-Factor | | Two-Factor | | Three-Factor | |
| Model | DOF | LR | DOF | LR | DOF | LR | DOF | LR |
| Two-Factor | | | | | | | | |
| TVMPR | 1 | 5.6043 | – | – | 1 | 7.7306 | – | – |
| | | (0.0179) | | – | | (0.0054) | | – |
| Three-Factor | | | | | | | | |
| Constant MPR | 5 | 1744.30 | – | – | 5 | 1209.00 | – | – |
| | | (0.0000) | | – | | (0.0000) | | – |
| TVMPR | 7 | 1754.53 | 2 | 10.2306 | 7 | 1215.40 | 2 | 6.3962 |
| | | (0.0000) | | (0.0060) | | (0.0000) | | (0.0408) |

### 3.5.2 Three-factor Model

If we compare the parameters with the three-factor model, the parameters of the long-term factor are quite similar. However, as both of the two short-term factors have significant mean-reversion and volatility coefficients, it suggests that the additional short-term factor is necessary for modelling crude oil prices. Also, the correlations between the three-factors are quite low and suggest that movements in the three-factors are largely independent of each other.

Investigating the improvement in model fit of the three-factor model over the two-factor model, it can be seen that the measurement errors are much lower for most contract maturities. The autocorrelation parameter of the measurement errors is also lower than the two-factor model, with a value of 0.87 although it still suggests some model misspecification. If we conduct an LR test with the two-factor constant

MPR against the three-factor models, then the likelihood ratio tests are for 5 and 7 parameter restrictions. The LR statistics in this case are 1744.30 and 1154.53 for the three-factor constant and time-varying MPR models respectively. These are much larger than the $\chi^2_{0.05,5}$ and $\chi^2_{0.05,7}$ values of 11.07 and 14.07 implying that the two factor model is highly misspecified. Similarly, if we consider the futures pricing errors for both the two-factor and three-factor models a large improvement in the cross-sectional fit can be observed. Figure 3.2 plots the difference between the observed and estimated log futures price for the 1 month, 1 year and 5 year contracts and both the two- and three-factor TVMPR models. The plot shows that the errors for the two-factor model are much more variable than the three-factor model and the magnitudes of the errors are also higher. This can be seen for each of the three contracts clearly showing that the three-factor model errors are lower.

Further evidence for the improvement in futures pricing can be found in Table 3.4. The table reports the mean error and the root-mean squared error (RMSE) of the estimated futures prices for each contract maturity both in-sample and out-of-sample under the two- and three-factor TVMPR models. The in-sample results are based on the estimated futures prices for the full-sample and the sub-sample period fom 1999-2006. The out-of-sample errors are obtained from the estimated futures prices in 2006-2008 using the parameters estimated in the sub-sample period of 1999-2006. Also, as the futures pricing fit is very similar for the constant and time-varying MPR versions of the models, only the time-varying versions were compared. The results show that the three-factor model outperforms the two-factor model both in-sample and out-of-sample for nearly every contract. Both the mean-error and the RMSE of the two-factor model are at least double the magnitude of the three-factor model for most contracts. The mean-errors of the two-factor models is negative for short and long-term contracts and positive for medium-term contracts which suggests that the two-factor model is unable to fit the entire term structure. Although both

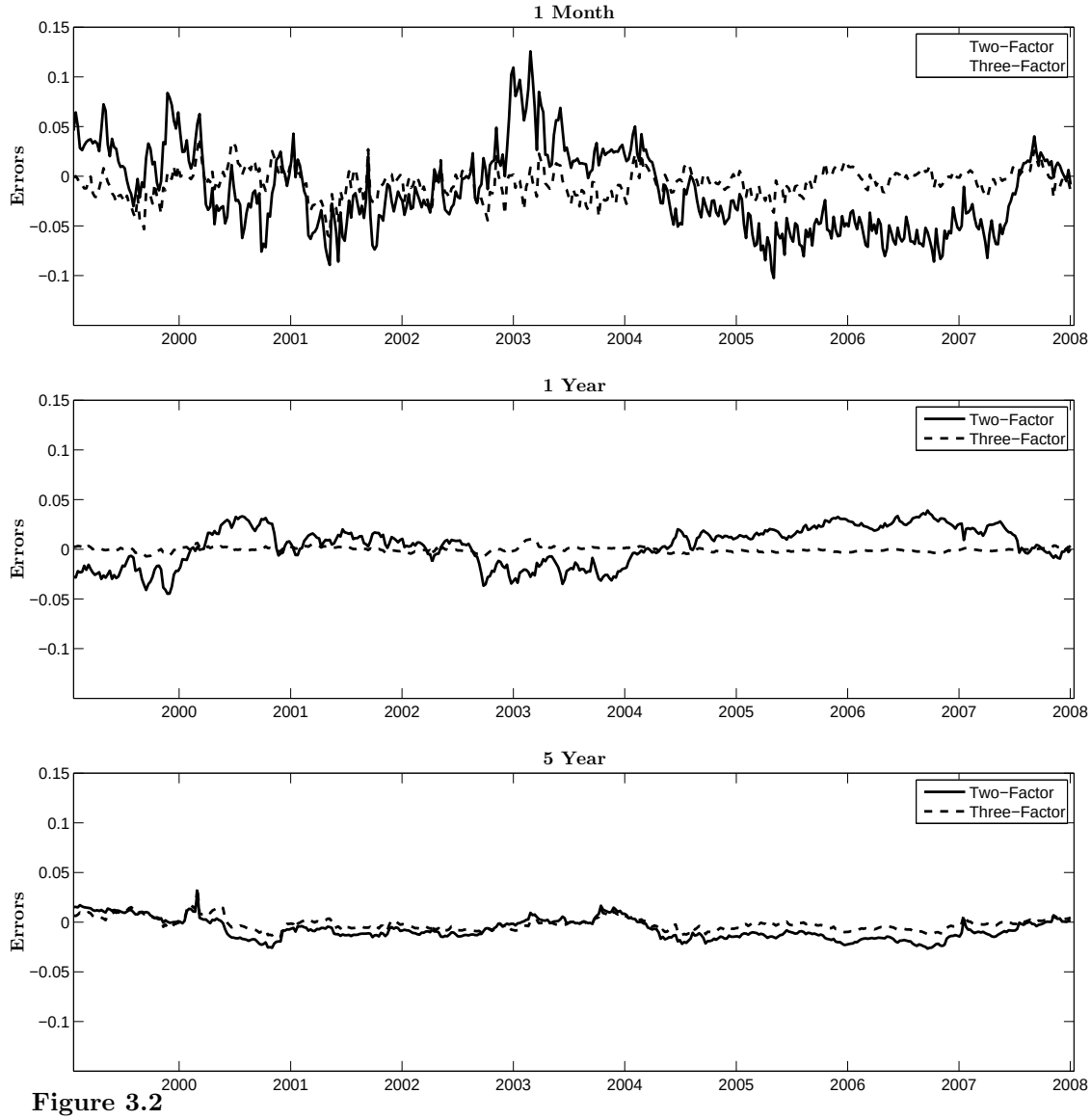



**Figure 3.2**
Plot of the difference between the true and estimated log futures price for 1-month, 1 year and 5 year contracts, Feb 1999 - Aug 2007

models tend to have higher errors for 1 month and 5 year contracts, the three-factor model exhibits much lower pricing errors across the full term structure with RMSE only exceeding 1% for 1 month contracts. These results hold both in-sample and out-of-sample which indicates that the two-factor model is largely misspecified and that three factors are necessary for modelling the term-structure of crude oil futures prices.

Looking at the MPR parameters implied by the three-factor model for the full-

**Table 3.4**
Futures Pricing Errors

This table reports the mean errors and root mean-squared errors (RMSE) of the difference between the observed log futures price and the estimated log futures price for each maturity. The in-sample results are the futures pricing errors for the full-sample and the sub-sample period, 1999-2006. The out-of-sample results are the futures pricing errors during 2006-2008 from the model estimated using data during 1999-2006. The errors are measured in percentage points.

| | In-Sample | | | | Out-of-Sample | |
|---|---|---|---|---|---|---|
| | Full Sample | | 1999-2006 | | 2006-2008 | |
| | Two-Factor | Three-Factor | Two-Factor | Three-Factor | Two-Factor | Three-Factor |
| Mean Error | | | | | | |
| 1 month | −1.51 | −0.66 | −1.01 | −0.70 | −3.83 | 0.06 |
| 3 months | −0.56 | −0.02 | −0.28 | −0.01 | −1.90 | 0.21 |
| 6 months | 0.00 | 0.13 | 0.00 | 0.18 | 0.00 | −0.06 |
| 1 year | 0.40 | −0.03 | 0.11 | −0.04 | 1.89 | −0.10 |
| 18 months | 0.55 | −0.14 | 0.22 | −0.19 | 2.42 | 0.06 |
| 2 years | 0.65 | −0.11 | 0.41 | −0.14 | 2.24 | 0.17 |
| 3 years | 0.48 | 0.02 | 0.39 | 0.01 | 1.27 | 0.26 |
| 4 years | −0.00 | 0.00 | −0.00 | −0.00 | −0.03 | −0.00 |
| 5 years | −0.71 | −0.19 | −0.70 | −0.21 | −1.34 | −0.48 |
| RMSE | | | | | | |
| 1 month | 4.36 | 1.60 | 4.21 | 1.66 | 4.97 | 0.80 |
| 3 months | 2.25 | 0.43 | 2.21 | 0.41 | 2.44 | 0.31 |
| 6 months | 0.00 | 0.29 | 0.00 | 0.35 | 0.00 | 0.16 |
| 1 year | 1.97 | 0.25 | 1.89 | 0.26 | 2.27 | 0.19 |
| 18 months | 2.53 | 0.36 | 2.42 | 0.45 | 2.95 | 0.21 |
| 2 years | 2.39 | 0.65 | 2.31 | 0.77 | 2.78 | 0.47 |
| 3 years | 1.20 | 0.47 | 1.10 | 0.51 | 1.57 | 0.49 |
| 4 years | 0.00 | 0.00 | 0.04 | 0.00 | 0.04 | 0.00 |
| 5 years | 1.25 | 0.67 | 1.27 | 0.74 | 1.60 | 0.72 |

sample, the results are fairly similar to the two-factor model. The values of $\alpha_1$, $\alpha_2$ and $\mu_3^*$ all suggest the market is generally in backwardation as the long-term means of long-term futures contracts tend to be lower than the long-term means of short-term contracts. In terms of the time-varying MPR parameters in both of the short-term factors, the constant risk-neutral parameters under a time-varying MPR specification are fairly similar to the constant MPR specification. However, only $\alpha_2$ and $\beta_2$ are significant at the 5% level which suggests that the time-varying MPR is mostly generated by only one of the short-term factors. However, as $\alpha_1$ is not significant for either the constant or time-varying MPR models, it could mean that

either more data is needed to accurately estimate both $\alpha_1$ and $\beta_1$ or a non-linear relationship may exist between the MPR and one of the short-term factors. As an additional test for time-varying MPR, an LR test is conducted with the restrictions that $\beta_1 = \beta_2 = 0$. The LR test results can be found in the "Three Factor" columns in Table 3.3 where the LR statistic is 10.23. This is larger than the $\chi^2_{0.05,2}$ critical value of 5.99 with a p-value of 0.006 so we can reject the hypothesis that the MPR is constant in the crude oil futures market. These results also hold in the sub-sample period.

In terms of the implications of the time-varying MPR, Casassus and Collin-Dufresne (2005) has suggested that a desirable property of commodity models is that we can distinguish between different sources of mean-reversion. One of the sources of mean-reversion is due to the negative correlation between risk premiums and spot prices. Evidence that confirms this in the crude oil market can be seen from the parameter estimates. In this case $\beta_2$ is negative and whilst $\beta_1$ is positive, it has a lower absolute value than $\beta_2$ and is not statistically significant. This suggests that negative correlation does exist between risk premiums (or the MPR) and spot prices. Hence in terms of our model, mean-reversion in spot prices can be seen to be generated by the negative correlation between risk premiums and spot prices and from short-term fluctuations in prices due to storage costs, transportation and financing costs as well as supply and demand factors.

### 3.5.3 Forecasting

Incorporating a time-varying MPR specification also has implications in terms of forecasting as the distributions of the underlying state variables differ under either constant or time-varying MPR. As a means of looking at whether the time-varying

MPR specification improves forecasting performance, the following approach is taken: At each time-step, given the model parameters and estimated state variable values, generate forecasts of the future spot price under the real-world measure at maturity where the maturity date is assumed to occur on the 1st day of the maturity month. The forecast error is then defined as the difference between the actual value at maturity of the contract and the expected spot price. This is given by

$$e_t = S_T - \mathbb{E}_{\mathbb{P}}(S_T|\mathcal{F}_t), \tag{3.15}$$

where $S_T$ is the spot price of the futures contract at $T$ and $e_t$ is the error at $T$ for the forecast made at time $t$. The spot price data is obtained from the U.S. Energy Information Administration[24] (EIA). To compare the errors generated from each of the model specifications, both the mean forecast errors and the RMSE is reported. The forecasts are conducted both in-sample and out-of-sample in a similar manner as for the futures pricing errors except both the constant and time-varying MPR models are included.

The in-sample forecasts are conducted for the full-sample and the sub-sample period of 1999-2006. The out-of-sample forecasts are conducted using the parameters estimated in the sub-sample period of 1999-2006 to compute the expected spot price during 2006-2008. Since only 2 years of data is used for out-of-sample forecasts, forecasts of up to 18 months only are included. Also, to benchmark the out-of-sample forecasting performance, the futures price and the spot price are included[25]. In these two cases, the forecasts are $\mathbb{E}_{\mathbb{P}}(S_T|\mathcal{F}_t) = F(t,T)$ for the futures price and $\mathbb{E}_{\mathbb{P}}(S_T|\mathcal{F}_t) = S_t$ for the spot price where the nearby futures price is used as a proxy

[24] http://www.eia.doe.gov

[25] Using the spot price as a forecast of future spot prices implies a random walk model. However, Alquist and Kilian (2010) found that it performed better than a number of econometric models for forecasting crude oil spot prices and is thus a suitable benchmark.

for the current spot price.

The forecasting results can be found in Table 3.5. Concentrating on the in-sample results first, all models seem to overestimate future spot prices as the mean errors are negative for all maturities in the full-sample and sub-sample. The forecasting errors tend to be quite high for all models and maturities with RMSE exceeding 20% for a number of maturities. However, the bias tends to be much lower for the time-varying MPR versions of the models as both the mean errors and RMSE are lower. In fact, it can be observed that at longer maturities, the two-factor time-varying MPR model outperforms the three-factor constant MPR model. This suggests that incorporating a time-varying MPR may be more important for forecasting than the addition of an extra factor.

For the out-of-sample forecasts, the results are consistent with the in-sample results. Of the two benchmarks, the futures price performed slightly better than the spot price for forecasting future oil prices. Between the estimated model specifications, the three-factor time-varying MPR also performed the best out-of-sample. However, none of the estimated models clearly outperformed the two benchmarks. The three-factor time-varying MPR model did perform better than futures prices for 1 and 3 month contracts but performed worse for 6, 12, and 18 month contracts. This does not provide any conclusive evidence that the short/long factor model performs better than the benchmark models in out-of-sample forecasts. However, between the different model specifications, the in-sample and out-of-sample forecasts does suggest that including a time-varying MPR specification improves forecasting performance.

Whilst the results are restricted to the sample studied, they show that a time-varying MPR specification can be important for modelling crude oil prices. This

**Table 3.5**

Forecast Errors

This table shows the in-sample mean and root mean-squared errors (RMSE) of the difference between the observed log spot price ($\log S_T$) at maturity and the forecast value ($\log \mathbb{E}_{\mathbb{P}}(S_T)$) for each model. The in-sample results are the forecast errors for the full-sample and the sub-sample period, 1999-2006. The out-of-sample results are the forecast errors during 2006-2008 from the model estimated using data during 1999-2006. The errors are measured in percentage points.

| Panel A: In-sample | Full Sample | | | | 1999-2006 | | | |
|---|---|---|---|---|---|---|---|---|
| | Two-Factor | | Three-Factor | | Two-Factor | | Three-Factor | |
| | Const | TV | Const | TV | Const | TV | Const | TV |
| Mean Error | | | | | | | | |
| 1 month | −3.87 | −3.78 | −2.39 | −2.26 | −4.01 | −3.90 | −3.31 | −3.12 |
| 3 months | −4.27 | −3.89 | −2.33 | −1.84 | −4.52 | −4.02 | −3.41 | −2.79 |
| 6 months | −6.37 | −5.09 | −3.78 | −2.37 | −6.34 | −4.81 | −4.71 | −3.01 |
| 1 year | −10.72 | −7.35 | −7.38 | −4.06 | −11.50 | −7.93 | −9.59 | −5.35 |
| 18 months | −14.32 | −9.47 | −10.55 | −5.81 | −17.37 | −12.29 | −15.63 | −9.12 |
| 2 years | −15.76 | −10.02 | −11.77 | −6.11 | −20.76 | −14.67 | −19.30 | −11.02 |
| 3 years | −20.40 | −13.98 | −16.40 | −9.81 | −26.61 | −19.37 | −25.40 | −15.17 |
| 4 years | −18.36 | −11.77 | −14.64 | −7.45 | −25.29 | −17.61 | −24.33 | −12.89 |
| 5 years | −19.86 | −13.15 | −16.23 | −8.45 | −27.38 | −19.44 | −26.35 | −13.96 |
| RMSE | | | | | | | | |
| 1 month | 9.96 | 9.76 | 8.66 | 8.54 | 10.32 | 9.86 | 9.38 | 9.23 |
| 3 months | 14.89 | 14.10 | 13.78 | 12.46 | 15.04 | 13.42 | 14.22 | 12.91 |
| 6 months | 19.09 | 18.15 | 18.39 | 15.76 | 18.30 | 16.19 | 17.80 | 15.08 |
| 1 year | 25.25 | 22.86 | 25.13 | 20.10 | 24.97 | 21.25 | 25.11 | 19.69 |
| 18 months | 28.26 | 25.15 | 27.98 | 22.29 | 30.34 | 26.40 | 30.60 | 24.11 |
| 2 years | 29.46 | 26.10 | 28.37 | 24.04 | 32.52 | 28.30 | 32.28 | 26.04 |
| 3 years | 31.94 | 28.32 | 29.60 | 26.47 | 36.21 | 31.34 | 35.37 | 28.88 |
| 4 years | 30.45 | 26.89 | 28.36 | 25.18 | 35.07 | 29.92 | 34.38 | 27.31 |
| 5 years | 30.10 | 26.22 | 27.82 | 24.21 | 35.53 | 29.89 | 34.71 | 26.64 |

| Panel B: Out-of-Sample | 2006-2008 | | | | | |
|---|---|---|---|---|---|---|
| | | | Two-Factor | | Three-Factor | |
| | Futures | Spot | Const | TV | Const | TV |
| Mean Error | | | | | | |
| 1 month | 3.05 | 3.05 | −6.13 | −8.14 | −1.77 | −2.02 |
| 3 months | 4.52 | 6.71 | −9.00 | −13.03 | −5.91 | −6.01 |
| 6 months | 3.36 | 8.38 | −17.86 | −23.34 | −16.85 | −15.59 |
| 1 year | 4.30 | 10.87 | −32.12 | −35.41 | −33.69 | −28.49 |
| 18 months | 18.16 | 22.84 | −33.62 | −33.16 | −35.58 | −27.43 |
| RMSE | | | | | | |
| 1 month | 8.79 | 8.79 | 9.87 | 11.99 | 6.68 | 6.43 |
| 3 months | 13.82 | 14.04 | 16.20 | 19.71 | 13.74 | 12.42 |
| 6 months | 19.02 | 20.76 | 25.96 | 30.52 | 24.96 | 22.50 |
| 1 year | 17.96 | 22.52 | 37.53 | 40.47 | 38.26 | 32.99 |
| 18 months | 21.05 | 25.18 | 34.95 | 34.89 | 36.05 | 28.25 |

could potentially alter decisions in risk management applications such as value at risk (VaR). For instance, if the different levels of mean reversion under the real-world and risk-neutral measures are not accounted for, it could lead to an overestimation of the VaR measure, thus tying up capital that could be used in other areas.

## 3.6 Time Varying Risk Premiums

Having estimated the model under a time varying MPR, we can estimate the risk premium using (3.7) and the output from the Kalman filter. As the parameter estimation results suggest the two-factor model is misspecified, we only consider the time-varying MPR of the three-factor model. As mentioned, the Kalman filter provides the minimum mean squared estimate of the state variables allowing us to estimate the MPR and risk premium implied by the model. A plot of the estimated risk premiums is shown in Figure 3.3. The figure shows that risk premiums are fairly stable over the period although large increases in risk premiums are generally consistent with major market events.

The most prominent features of the plot include the low risk premiums exhibited during 2000 when oil prices tripled between January 2000 and September 2000, a sharp increase in risk premiums following September 11 compounded by Enron filing for bankruptcy and the Argentinian government's default in December 2001, and the increase in risk premiums from late 2002 to early 2003 coinciding with the uncertainty in oil prices surrounding the Iraq invasion. From 2004 - 2007, risk premiums have tended to be relatively lower reflecting a period where oil prices were increasing rapidly as a result of a number of factors including supply disruptions due to political instability, an increase in world oil demand and constrained OPEC

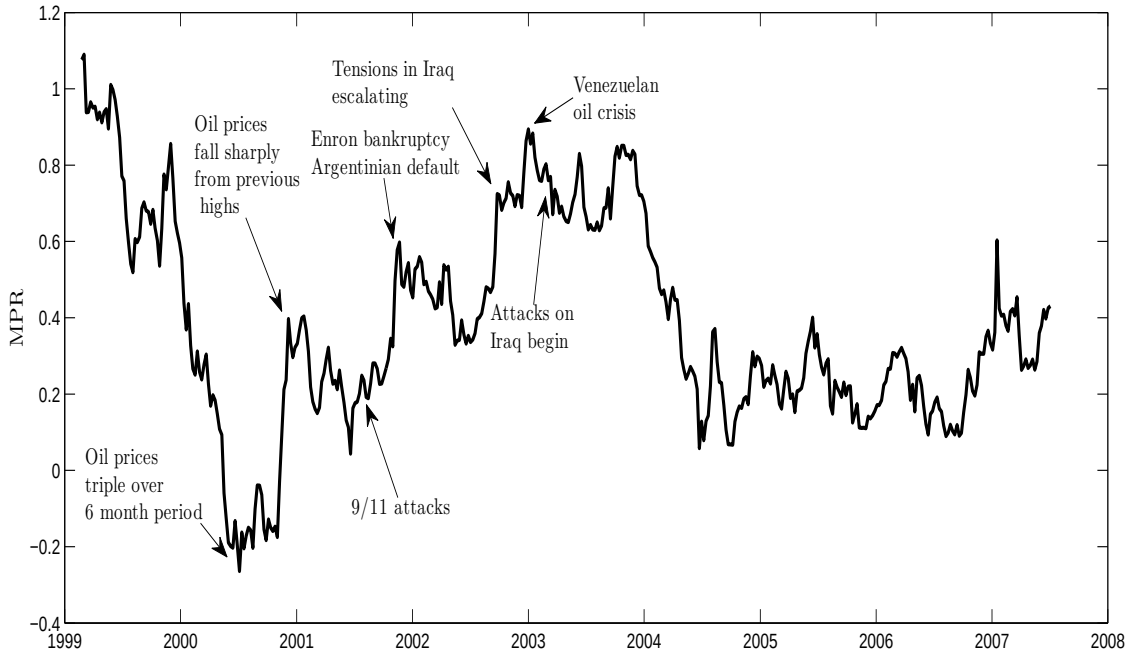



**Figure 3.3**
Plot of the market price of risk under the three-factor model

output. The lower risk premiums during this period are possibly an indication of the perception that oil is an increasingly scarce resource. As explained in Section 3.3, risk premiums can be thought of as insurance paid by sellers to speculators to insure them against a future fall in prices. Hence, in a market where the resource is scarce and future price rises are likely, risk premiums are likely to be lower. This period was also characterized by low historical volatility and credit spreads and high growth in asset prices worldwide. Hence risk premiums in equity and bond markets were most likely quite low during this period similar to the behaviour in the crude oil futures market.

The plot therefore highlights that although major changes in risk premiums are due to events idiosyncratic to crude oil markets, they are also consistent with systematic market events. This gives some indication that systematic risk factors may play a role in risk premiums exhibited in the crude oil market. Using the estimate of the risk premium, we can determine whether any relationship exists between the risk premiums in the crude oil market and equity and bond markets. As risk premiums

are unobservable, we consider the approach taken by Bailey and Chan (1993) and Baum and Barkoulas (1996) to determine whether factors that are indicative of risk premiums in stock and bond markets give similar information in the crude oil market. The factors considered in Bailey and Chan (1993) and Baum and Barkoulas (1996) are credit spreads, dividend yields and treasury bond yields. Keim and Stambaugh (1986) and Fama and French (1989) found that credit spreads and dividend yields have forecast power for the returns on stocks and bonds leading Bailey and Chan (1993) to suggest that this allows them to be considered as indicators of risk premiums in equity and bond markets. Furthermore, treasury bill yields give an indication of the state of the economy which tends to be lower during economic downturns and higher during growth periods. As an additional indicator of market risk, this study also considers the market return variance or volatility. In this case, the CBOE volatility index (VIX) is considered. The VIX index is a model-free measure of implied volatility of S&P500 index options that represents the market's expectation of volatility over the next 30 days. As such, the VIX index is able to give an indication of the expected risk in equity markets. If risk premiums in the crude oil market contain systematic risk factors, then these measures should be able to explain the variation in the market price of risk.

To test whether these macroeconomic risk factors provide information about crude oil risk premiums, we conduct a linear regression of the crude oil risk premium and the variables which represent risk premiums in equity and bond markets. The multivariate regression includes the four variables above - credit spreads, dividend yields, treasury bond yields and market volatility. The regression equation is as follows,

$$\Lambda_t = a_0 + a_1 \text{CRSPR}_t + a_2 \text{SPDY}_t + a_3 \text{TB}_t + a_4 \text{VIX}_t + \epsilon_t, \tag{3.16}$$

where $\Lambda_t$ is the risk premium defined by (3.7).

The data used covers the same sample period as the crude oil futures data consisting of weekly observations spanning the period from Feb 1999 - Aug 2007. The variables are defined as follows, the credit spread (CRSPR) is the difference between the yield on 30 year Baa and Aaa Moody's rated bonds, dividend yield (DY) is the dividend yield on the S&P500 index, the treasury bond (TB) yield is the yield on the 3-month constant maturity US Treasury bill and VIX is the Chicago Board Options Exchange Volatility Index. The corporate and treasury bond yield data is obtained from the Board of Governors of the Federal Reserve System website whilst the dividend yield and VIX data were obtained from Datastream. If we consider all of the variables of interest a priori, credit spreads, dividend yields and market volatility should be positively related to risk premiums whilst interest rates should be negatively related to risk premiums in equity and bond markets.

**Table 3.6**
Regression of the market price of risk and macroeconomic factors

This table gives the parameter estimates for the regression equation given by (3.6) obtained by OLS. The Newey and West (1987) heteroscedasticity and autocorrelation consistent standard errors are in parentheses.

| Parameter | Constant | CRSPR | SPDY | TB | VIX | $\bar{R}^2$ |
|---|---|---|---|---|---|---|
| Estimate | 2.9127 | 26.5735 | 2.1310 | -3.3062 | 1.0323 | 0.2130 |
| S.E. | (26.4323) | (12.9079) | (11.5379) | (1.6732) | (0.4467) | |

The results of the regression can be found in Table 3.6. Credit spreads and the VIX index have significantly positive relationships with the risk premium, whilst treasury bill yields have a significantly negative relationship. Although dividend yields were found to have no significant relationship with the MPR this does not suggest that equity risk premiums play no role in crude oil risk premiums. It could suggest either that equity volatility risk plays a larger role in crude oil risk premiums than equity risk premiums or that dividend yields provide less information about equity risk

premiums than volatility. Overall, these results are broadly what would be expected as large credit spreads, high volatility and low interest rates are indicative of high risk premiums in equity and bond markets and these are able to explain some of the risk premiums in the crude oil futures market. If we also consider the adjusted $R^2$ value of 0.213, it implies that risk premiums in the crude oil futures market can be partially explained by risk factors common to equity and bond markets. The results agree with some of the previous findings for other non-energy related commodities, demonstrating that the perception of risk in crude oil markets is partially due to systematic risk factors in equity and bond markets.

## 3.7 Conclusion

This chapter estimates a three-factor model for crude oil that incorporates a time-varying market price of risk specification. The model considered is an extension of the short/long term factor model of SS which incorporates both an additional short-term factor and a time-varying MPR. The model is intuitively appealing as it can be interpreted as consisting of two components - a long-term component representing an equilibrium price level which is uncertain and a short-term component consisting of two mean-reverting factors representing the short-term variations in prices from the equilibrium price level. Also, the MPR is specified as a linear function of the two short-term factors allowing risk premiums to be time-varying.

Using the Kalman filter to estimate the model parameters, it is found that the model captures the term-structure of futures prices more accurately than the two-factor model. Additionally, by incorporating the time-varying MPR specification the model is better able to describe the time-series behaviour of crude oil spot prices

where it was found that the spot price has a different level of mean-reversion under the real-world measure compared to the mean-reversion implied by futures prices under the risk-neutral measure. Whilst the model fit supports the model and MPR specification, like most other approaches to modelling asset prices, the choice of MPR specification is arbitrary. By extracting the MPR and risk premium implied by the model parameters, it allows us to further investigate the MPR specification and whether it is economically justified.

The study examines the relationship between risk premiums in the crude oil futures market and risk factors in the macroeconomy. Our findings confirm that whilst the main determinants of risk premiums are idiosyncratic, macroeconomic risk factors which drive risk premiums in equity and bond markets have explanatory power for risk premiums in the crude oil market. The results suggest that incorporating a time-varying MPR is important when modelling crude oil futures prices as they are correlated with systematic risk factors in equity and bond markets. This leads to possible future avenues of research in the area such as an investigation of modelling approaches which incorporate systematic risk factors directly in the model dynamics. The findings could also have implications for portfolio management as using crude oil futures may be useful in hedging and managing portfolio risk. Finally, the modelling approach is not restricted to crude oil or energy-related products and incorporating a time-varying MPR specification may be suitable for other commodities which exhibit similar characteristics.

# Chapter 4

# Commodity Modelling with Stochastic Volatility and Jumps

In the previous chapter, a three-factor short/long factor model was used to estimate futures prices in the crude oil futures market. However, the model is Gaussian and in light of recent history (the most recent being the global financial crisis), the model is not able to explain some of the characteristics exhibited by commodity prices. Although the model can account for mean-reversion in short term prices, expected increases in long-term prices and the "Samuelson Effect", it fails to account for non-Gaussian behaviour. This includes large changes in spot prices, time-varying volatility and volatility clustering. In particular, evidence for the latter two effects were found in Chapter 2, where a discrete-time stochastic volatility model was examined for crude oil volatility. The study showed that time-varying volatility is present in crude oil prices and that volatility is highly persistent which is evidence of the volatility clustering effect[26]. However, the discrete-time approach does not easily facilitate pricing derivatives and neither does the form of the stochastic volatility factor. Alternatively, continuous-time models and affine jump-diffusion models in particular, are appealing for pricing derivatives as they allow the use of established pricing frameworks under the assumption of no-arbitrage.

[26]The volatility persistence parameter for crude oil in Chapter 2 was found to be above 0.95. This implies that volatility in each period will be highly correlated with the previous period's volatility level and indicates volatility clustering is present.

The commodity modelling literature has considered continuous-time models with non-Gaussian dynamics where a number of authors have introduced jumps and stochastic volatility to model commodity prices. However, this is mainly done separately and it is only recently that stochastic volatility has gained some attention. On the other hand, introducing jumps in model dynamics has been somewhat more popular in the literature. Some authors who have taken this approach include Villaplana (2002), Casassus and Collin-Dufresne (2005), Aiube et al. (2008), Askari and Krichene (2008), Dempster et al. (2008) and Crosby (2008) among others. The modelling approach of these authors fall under the affine jump-diffusion class of models which allows for tractable solutions to derivative prices. These studies take either a single or multi-factor approach (usually up to a maximum of three factors) to model spot prices, where the empirical studies were mainly applied to futures and/or the physical (spot) markets.

In terms of stochastic volatility modelling, Richter and Sørensen (2002), Hikspoors and Jaimungal (2008), Hughen (2010) and Trolle and Schwartz (2009) incorporate stochastic volatility in the model dynamics to price commodities. Of these studies, Richter and Sørensen (2002), Hikspoors and Jaimungal (2008), Hughen (2010) proposed models which explain the spot price dynamics using an affine set of state variables including a stochastic volatility factor. Trolle and Schwartz (2009) apply the Heath, Jarrow and Morton (1992) (HJM) approach to commodity futures pricing which incorporates stochastic volatility. However, the commodity literature has not paid much attention to stochastic volatility modelling where it has been relatively unpopular until recently.

For most of the studies on commodity markets, the focus has been on futures markets. The commodity pricing literature also tends to be quite sparse when it comes to models with stochastic volatility or jumps in price dynamics. Compared to the

financial market literature, it is much smaller and few have considered options pricing performance. Of the studies above, only Richter and Sørensen (2002), Hughen (2010) and Trolle and Schwartz (2009) analysed option pricing empirically. For some of the studies on jump-diffusion models, only theoretical option prices were considered. Also, the only study that we are aware of which considers a commodity model with both jumps and stochastic volatility to price commodity derivatives is a paper by Yan (2002). However, although the paper derives futures and option prices, it does not provide an empirical analysis of the model and so the model's performance for pricing commodities is not known. A likely reason for the lack of empirical studies on pricing commodity options could be due to data being more easily accessible in equity and bond markets than commodity markets. Those markets are also larger and have higher liquidity.

However, the market for other derivatives has grown considerably, especially for crude oil, as options have become increasingly popular for hedging and risk management as an alternative to futures contracts. Recent reports by the U.S. Commodity Futures Trading Commission (CFTC) found that at the start of 2000, the proportion of the open interest in delta-adjusted WTI crude oil options[27] to the total open interest in futures and options traded on NYMEX amounted to around 27%. However, by the start of 2009, the proportion increased to about 60% of total open-interest. This means that a majority of derivatives traded on NYMEX are now options rather than futures. Given the increased importance of options, it is necessary that models can price options accurately.

In this chapter, the three-factor short/long factor model is extended to incorporate both stochastic volatility and jumps. The extension follows similar approaches in

[27] "Option open interest and traders' option positions are computed on a futures-equivalent basis using delta factors supplied by the exchanges. Long-call and short-put open interest are converted to long futures-equivalent open interest. Likewise, short-call and long-put open interest are converted to short futures-equivalent open interest." Source: http://www.cftc.gov.

equity markets such as Heston (1993) who introduced stochastic volatility to the Black-Scholes model and Bates (1996) who introduced jumps to the Heston (1993) model. For the volatility process, we follow Heston (1993) and introduce a square-root volatility process. For the jumps, as the model consists of three spot price factors, jumps are introduced for each factor. The jump specification allows for some jumps to affect short-term futures prices more than long-term futures prices, a feature deemed necessary for commodities by Crosby (2008). The approach follows Dempster et al. (2008) which nests a number of jump-diffusion models introduced in the commodity market literature, although their approach does not consider stochastic volatility. The model also allows for a time-varying market price of risk specification in both the spot price and volatility processes. This chapter adds to the literature by investigating commodity models with stochastic volatility and jumps. The study is conducted empirically and analyses the futures and option pricing performance. As in previous chapters, although the model can be applied to other commodities, the empirical study is restricted to the crude oil futures market as it is the most liquid commodity market. This study attempts to fill a gap in the literature by introducing a commodity model with both jumps and stochastic volatility and empirically analysing the model using futures and options data.

In this study, we derive semi-closed form solutions to futures and options prices for the stochastic volatility jump-diffusion model under the assumption of no-arbitrage. We empirically test the impact of introducing either jumps or stochastic volatility for estimating futures prices and compare the results to the Gaussian three-factor model introduced in Chapter 3. The models are estimated using the estimation approaches analysed in Chapter 2. In particular, the jump-diffusion model is estimated using a Monte Carlo likelihood (MCL) approach whereas the SV model is estimated using a quasi-maximum likelihood (QML) approach. The findings suggest that although there is little improvement in futures pricing performance, there is evidence that the

model is able to explain the distributional properties of crude oil prices better when either jumps or stochastic volatility are added. The option pricing performance of the models are also compared. This is initially conducted by pricing European call and put options using model parameters estimated from futures prices only.

As the pricing performance was found to be quite poor for all models when options are priced this way, a more direct approach of estimating option prices is considered. This involves estimating the model parameters using both futures and option prices in the estimation process. As there is a nonlinear relationship between option prices and state variables, the linear Kalman filter is inappropriate for estimation. Instead, we use the Unscented Kalman Filter (UKF) to estimate the model parameters. We estimate the fully specified stochastic volatility jump-diffusion model, which enables a comparison of its futures and options pricing performance to the nested specifications.

This chapter is organised as follows - Section 4.1 describes the extension which incorporates stochastic volatility in the model, Section 4.2 derives pricing equations for futures and European futures options for the short/long SVJ model, Section 4.3 presents the empirical results using futures prices to estimate the models, Section 4.4 presents the empirical results using both futures and option prices to estimate the models and 4.5 concludes the chapter.

## 4.1 Short/Long Factor Stochastic Volatility Jump-Diffusion Model

We assume a fixed probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and the filtration $\mathcal{F} = (\mathcal{F}_t)_{t\geq 0}$. The three-factor long/short model of the previous chapter can be extended to incorporate stochastic volatility and jumps as follows,

$$\begin{aligned} dX_{1,t} &= -\kappa_1 X_{1,t}dt + \sigma_1\sqrt{V_t}dZ_{1,t} + J_{1,t}dN_{1,t}, \\ dX_{2,t} &= -\kappa_2 X_{2,t}dt + \sigma_2\sqrt{V_t}dZ_{2,t} + J_{2,t}dN_{2,t}, \\ dX_{3,t} &= \mu_3 dt + \sigma_3\sqrt{V_t}dZ_{3,t} + J_{3,t}dN_{3,t}, \\ dV_t &= \kappa_v(\theta_v - V_t)dt + \sigma_v\sqrt{V_t}dZ_{v,t}, \end{aligned} \tag{4.1}$$

where $Z_{i,t}$, $i = 1, 2, 3, v$ are correlated Brownian motions with instantaneous correlation coefficients $dZ_{i,t}dZ_{j,t} = \rho_{ij}dt$, $dN_{i,t}$ are Poisson processes with jump intensity $\mathbb{P}(dN_{i,t} = 1) = \lambda_i dt$ and jump size density $f_{J_i}(j)$. The jumps have the following distributions - the jumps in the first factor are positive exponentially distributed, the jumps in the second factor are negative exponentially distributed and the jumps in the third factor are Gaussian, i.e.

$$J_{1,t} \sim \exp(\omega_u),\ -J_{2,t} \sim \exp(\omega_d) \text{ and } J_{3,t} \sim N(0, \sigma_j^2). \tag{4.2}$$

The volatility factor, $V_t$, follows the well-known square-root process which has the desirable property of being non-negative and never hits the zero boundary under certain parameter restrictions. This condition is well-known in the literature and for this model, the parameter restriction is

$$2\kappa_v\theta_v > \sigma_v^2.$$

It is assumed throughout this chapter that this condition holds. Additionally, the three spot price factors are governed by the volatility factor although the degrees to which they depend on the volatility factor is determined by $\sigma_1$, $\sigma_2$ and $\sigma_3$.

In the empirical application, three nested models are considered. They are a three-factor jump-diffusion model with no stochastic volatility component ($V_t$ is constant), a stochastic volatility model with no jumps, (i.e. $\lambda_1 = \lambda_2 = \lambda_3 = 0$) and a three-factor Gaussian model with no stochastic volatility or jumps. We will refer to the jump-diffusion model as JD, the stochastic volatility model as SV and the Gaussian model as Gaussian. Also, the stochastic volatility jump-diffusion model will be referred to as SVJ. In the analysis of the models in later sections, we will refer to $X_{1,t}$, $X_{2,t}$ and $X_{3,t}$ as the spot price factors and $V_t$ as the volatility factor.

As the model given by (4.1) is specified under the real-world measure, a specification of the risk-neutral measure is needed to derive derivatives prices. In order to achieve this, we make the usual assumption that the market admits no arbitrage. However, when stochastic volatility and jumps are introduced, the market is incomplete. This means that when specifying the risk-neutral measure, there is no unique choice for the market price of risk specification. In order to obtain one, the approach taken here is to assume that the market price of risk (MPR) is derived from the market although an assumption is made about the functional form of the MPR. To facilitate application of Girsanov's theorem and specifying the market price of risk, the model is reformulated in terms of standard Brownian motion. The model given by (4.1) can be equivalently represented as

$$dX_t = (\mu + \kappa X_t)dt + \sqrt{V_t}\Sigma dW_t, \tag{4.3}$$

where $X_t = (X_{1,t}, X_{2,t}, X_{3,t}, X_{v,t})'$, $X_{v,t} \equiv V_t$, $\mu = (0, 0, \mu_3, \kappa_v\theta_v)'$,

$$\kappa = \begin{pmatrix} -\kappa_1 & 0 & 0 & 0 \\ 0 & -\kappa_2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -\kappa_v \end{pmatrix}, \; \Sigma\Sigma' = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} & \sigma_{1v} \\ \sigma_{12} & \sigma_{22} & \sigma_{23} & \sigma_{2v} \\ \sigma_{13} & \sigma_{23} & \sigma_{33} & \sigma_{3v} \\ \sigma_{1v} & \sigma_{2v} & \sigma_{3v} & \sigma_{vv} \end{pmatrix},$$

$\Sigma\Sigma'$ positive semidefinite with $\sigma_{ij} = \rho_{ij}\sigma_i\sigma_j$, $i \neq j$, $\sigma_{ij} = \sigma_i^2$, $i = j$ for $i = 1, 2, 3, v$, and $W_t = (W_{1,t}, W_{2,t}, W_{3,t}, W_{v,t})'$ is 4-dimensional standard Brownian motion.

As the estimation approach requires pricing derivatives under the risk-neutral measure and estimation of the underlying state variables, the change of measure must be specified. The existence of an equivalent martingale measure (EMM) under Heston stochastic volatility has been established in the literature when certain parameter restrictions apply. However, as the market is incomplete under the model when either stochastic volatility or jumps are present, we make a few assumptions regarding the MPR. In terms of the jump component, a number of authors in the commodity pricing literature assume that the jump risk is diversifiable (e.g. Casassus and Collin-Dufresne (2005), Cartea and Figueroa (2008), Aiube et al. (2008) and Dempster et al. (2008)). This assumption follows from Merton (1976) who argued that jump risk in equity markets can be considered a non-systematic risk and is therefore diversifiable. Although a major reason for this assumption is for computational tractability when estimating the model parameters, a few of these studies (see Casassus and Collin-Dufresne (2005) and Dempster et al. (2008)) found little additional improvement in model fit when the jump intensity is allowed to change under a change of measure. Given these studies and to reduce the complexity of estimating the model, the same assumption is made here. This means that only the drift is adjusted in relation to the jump component when changing measures.

Although this assumption restricts the set of valid equivalent martingale measures, there is still scope for flexibility in the choice of an EMM due to the diffusion component. In the previous chapter, the MPR was allowed to be a linear function of each state variable. However, when stochastic volatility is present, the existence of an EMM is not guaranteed. Wong and Heyde (2006) shows that the EMM exists under the Heston (1993) model when the volatility process is strictly positive (i.e. the restriction $2\kappa_v\theta > \sigma_v^2$ is imposed). Cheridito et al. (2007) also demonstrates the existence of an EMM for term structure models with multi-dimensional square-root processes and an MPR which is affine in the state variables. They refer to their MPR specification as an "extended affine" market price of risk[28]. The only restriction under an "extended affine" MPR specification relates to the existence and boundary non-attainment conditions of square-root processes. Here the MPR is restricted so that the boundary non-attainment conditions are satisfied and each of the state factors remains structurally of the same class under the change of measure[29]. It is therefore possible to establish existence of the EMM.

We define the change of measure by the Radon-Nikodym derivative,

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = \exp\Big(-\int_0^T \gamma_t \cdot dW_t - \frac{1}{2}|\gamma_t|^2 dt\Big) = \eta_T, \tag{4.4}$$

[28]This is in contrast to Dai and Singleton (2000) who introduced the completely affine and Duffee (2002) who introduced the essentially affine MPR specifications. In terms of the relationship between the classes, the essentially affine class nests the completely affine class, whilst the extended affine class nests the essentially affine class.

[29]In this case, it means that only the MPR for $X_{1,t}$, $X_{2,t}$ and $V_t$ are allowed to be linear functions of themselves. The MPR for $X_{3,t}$ is assumed to be constant.

where $\gamma_t = (\sqrt{V_t}\Sigma)^{-1}(\alpha + \beta X_t + \lambda(\mathbb{E}_{\mathbb{P}}[e^{J_t}] - 1))$[30] and

$$\alpha = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_v \end{pmatrix}, \ \beta = \begin{pmatrix} \beta_1 & 0 & 0 & 0 \\ 0 & \beta_2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \beta_v \end{pmatrix} \text{ and } \lambda = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ 0 \end{pmatrix}.$$

When $\alpha_v = 0$, Wong and Heyde (2006) show that the EMM exists and the model admits no arbitrage. However, when $\alpha_v \neq 0$, it is not clear whether the Novikov or Kazamaki condition is satisfied and existence of the EMM can not be established this way. Alternatively, Cheridito et al. (2007) show that a requirement for existence of $\mathbb{Q}$ is that the zero-boundary for the volatility factor is never hit under either $\mathbb{P}$ or $\mathbb{Q}$[31]. Although this approach is applied to term-structure modelling, the same approach can be used here as jump risk is ignored which means that the MPR specification in the SVJ model is a special case of the extended affine specification. This leads to the following restriction on the volatility factor,

$$2\kappa_v^*\theta_v^* > \sigma_v^2 \Leftrightarrow \alpha_v < \kappa_v\theta_v - \frac{\sigma_v^2}{2}.$$

From Theorem 1 in Cheridito et al. (2007), given that the boundary non-attainment condition is satisfied, then $\mathbb{E}_{\mathbb{P}}(\eta_T) = 1$. Hence, by Girsanov's theorem, $dW_t^* = dW_t + \gamma_t dt$ is a 4-dimensional standard Brownian motion under $\mathbb{Q}$. Under the risk-neutral measure $\mathbb{Q}$, the model is specified by the following SDE,

$$dX_t = (\mu^* + \kappa^* X_t)dt + \sqrt{V_t}\Sigma dW_t^* + J_t dN_t, \tag{4.5}$$

[30] The additional term in $\gamma_t$, $\lambda(\mathbb{E}_{\mathbb{P}}[e^{J_t}] - 1)$ relates to the jump risk premium. This allows the jump terms to be martingales under $\mathbb{Q}$ with respect to the spot price and is added for convenience when computing derivative prices. As the terms are constant and only affect the drift of the spot price factors, it has no impact on the existence of the change of measure.

[31] As stated in Cheridito et al. (2007) in the context of term-structure models, by allowing $\alpha_v$ to be non-zero, this allows the market price of risk to be non-zero when $V_t$ is near zero. Previous literature did not make it clear whether arbitrage opportunities exist in this case. However, they showed that there is no arbitrage.

where

$$\mu^* = \mu - \alpha = \begin{pmatrix} -\alpha_1 - \lambda_1(\mathbb{E}_{\mathbb{P}}[e^{J_{1,t}}] - 1) \\ -\alpha_2 - \lambda_2(\mathbb{E}_{\mathbb{P}}[e^{J_{2,t}}] - 1) \\ \mu_3 - \alpha_3 - \lambda_3(\mathbb{E}_{\mathbb{P}}[e^{J_{3,t}}] - 1) \\ \kappa_v\theta_v - \alpha_v \end{pmatrix},$$

$$\kappa^* = \kappa - \beta = \begin{pmatrix} -(\kappa_1 + \beta_1) & 0 & 0 & 0 \\ 0 & -(\kappa_2 + \beta_2) & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -(\kappa_v + \beta_v) \end{pmatrix},$$

$$J_t = \begin{pmatrix} J_{1,t} & 0 & 0 & 0 \\ 0 & J_{2,t} & 0 & 0 \\ 0 & 0 & J_{3,t} & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

$\Sigma$ as defined in (4.3), $\beta_i < \kappa_i$ for $i = 1, 2, v$ and $N_t = (N_{1,t}, N_{2,t}, N_{3,t}, N_{4,t})'$ is a 4-dimensional Poisson process. Also, in terms of the nested MPR specifications, the completely affine specification is obtained when $\beta_1 = 0$, $\beta_2 = 0$ and $\alpha_v = 0$ and the essentially affine specification is obtained when $\alpha_v = 0$.

Transforming back to the specification with correlated Brownian motions,

$$\begin{aligned} dX_{1,t} &= (-\alpha_1 - \kappa_1^* X_{1,t})dt + \sigma_1\sqrt{V_t}dZ_{1,t}^* + J_{1,t}dN_{1,t} - \lambda_1(\nu_1(1) - 1)dt, \\ dX_{2,t} &= (-\alpha_2 - \kappa_2^* X_{2,t})dt + \sigma_2\sqrt{V_t}dZ_{2,t}^* + J_{2,t}dN_{2,t} - \lambda_2(\nu_2(1) - 1)dt, \\ dX_{3,t} &= \mu_3^* dt + \sigma_3\sqrt{V_t}dZ_{3,t}^* + J_{3,t}dN_{3,t} - \lambda_3(\nu_3(1) - 1)dt, \\ dV_t &= \kappa_v^*(\theta_v^* - V_t)dt + \sigma_v\sqrt{V_t}dZ_{v,t}^*, \end{aligned} \tag{4.6}$$

where $\kappa_i^* = \kappa_i + \beta_i$, for $i = 1, 2, v$, $\mu_3^* = \mu_3 - \alpha_3$, $\theta_v^* = (\kappa_v\theta_v - \alpha_v)/\kappa_v^*$, $Z_{i,t}^*$, $i = 1, 2, 3, v$ are correlated Brownian motions under the risk neutral measure $\mathbb{Q}$,

with instantaneous correlation coefficients $dZ^*_{i,t}dZ^*_{j,t} = \rho_{ij}dt$, for $i = 1, 2, 3, v$ and $i \neq j$ and $\nu_i(c) = \mathbb{E}_{\mathbb{Q}}[e^{cJ_{i,t}}]$. For the distribution of the jump sizes defined above,

$$\nu_1(c) = \frac{\omega_u}{\omega_u - c}, \tag{4.7a}$$

$$\nu_2(c) = \frac{\omega_d}{\omega_d + c}, \tag{4.7b}$$

$$\nu_3(c) = e^{c^2\sigma_j^2/2}. \tag{4.7c}$$

## 4.2 Pricing Derivatives

### 4.2.1 Futures Prices

Given the above risk-neutral dynamics of the model we can price a number of derivatives with semi-analytical formulas. For commodity markets, the main traded assets are futures contracts. For a futures contract which matures at time $T$, it is known that the futures price is the expected future spot price under the risk neutral measure (Cox et al. (1981)), i.e.

$$F(t,T) = E_{\mathbb{Q}}\big[S_T|\mathcal{F}_t\big], \tag{4.8}$$

where $\mathcal{F}_t$ is the information set at time $t$. As the model above is an exponentially affine jump-diffusion model, the solution to the futures price has the following form[32],

$$F(t,T) = \exp\big(A(\tau) + B(\tau)\cdot X_t\big), \tag{4.9}$$

[32]For instance, see Duffie et al. (2000)

where $\tau = T - t$ and $B(\tau) = (B_1(\tau), B_2(\tau), B_3(\tau), B_v(\tau))'$. In order to determine the solution to the futures price, we first state the generalized Itô's lemma for multidimensional jump-diffusion processes.

Let $X_t = (X_{1,t}, ..., X_{d,t})$ be a d-dimensional jump-diffusion process defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and is given by the following equation

$$dX_t = \mu(t, X_t)dt + \sigma(t, X_t)dW_t + J(t, X_t)dN_t, \tag{4.10}$$

where $\mu : [0, T] \times \mathbb{R}^d \to \mathbb{R}^d$ vector, $\sigma : [0, T] \times \mathbb{R}^d \to \mathbb{R}^{d\times d}$, $W_t$ is d-dimensional standard Brownian motion and $N_t$ is a d-dimensional Poisson process with jump sizes $J(t, X_t)$. Then for any $C^{1,2}$ function $f : [0, T] \times \mathbb{R}^d \to \mathbb{R}$, the process $Y_t = f(t, X_t)$ can be represented as,

$$\begin{aligned} dY_t =& \frac{\partial f}{\partial t}(t, X_t)dt + \sum_{i=1}^{d} \frac{\partial f}{\partial X_i}(t, X_t)\mu_i(t, X_t)dt + \sum_{i=1}^{d} \frac{\partial f}{\partial X_i}(t, X_t)\sigma_i(t, X_t)dW_{i,t} \\ &+ \sum_{i,j=1}^{d} \frac{1}{2}\frac{\partial^2 f}{\partial X_i \partial X_j}(t, X_t)\sigma_i(t, X_t)\sigma_j(t, X_t)'dt \\ &+ \sum_{i=1}^{d} \Big[ f(t, X_{t-} + \Delta X_{i,t}) - f(t, X_{t-}) \Big] dN_{i,t}, \end{aligned} \tag{4.11}$$

where $\mu_i(t, X_t)$ is the i$^{\text{th}}$ element of $\mu(t, X_t)$ and $\sigma_i(t, X_t)$ is the i$^{\text{th}}$ row of $\sigma_i(t, X_t)$.

Let $F_t \equiv F(t, T)$. Applying the generalized Itô's lemma to the futures price in (4.9) and using the SDE in (4.5) gives

$$\begin{aligned} dF_t = &\left[ \frac{\partial F}{\partial t} + \sum_{i=1,2,3,v} \frac{\partial F}{\partial X_i}(\mu_i^* + \kappa_i^* X_{i,t}) + \frac{1}{2} \sum_{i,j=1,2,3,v} \frac{\partial^2 F}{\partial X_i \partial X_j} V_t \sigma_{ij} \right] dt \\ &+ \sum_{i=1,2,3,v} \frac{\partial F}{\partial X_i}\sigma_i dZ_{i,t}^* + \sum_{i=1,2,3} F_{t-} \Big( \exp\{B_i(\tau)J_{i,t}\} - 1 \Big) dN_{i,t}, \end{aligned} \tag{4.12}$$

where $\mu_i^*$ denotes the i$^{\text{th}}$ element of $\mu^*$ in (4.5).

Since futures contracts cost nothing to enter, to preclude arbitrage opportunities, the futures price must be a martingale under the risk neutral measure. Noting that

$$\int_0^t \Big[\exp\{B_i(\tau)J_{i,s}\} - 1\Big] dN_{i,s} - \int_0^t \lambda_i \Big[\nu_i\big(B_i(\tau)\big) - 1\Big] ds,$$

is a martingale and setting the drift of the futures price to zero in (4.12) leads to the following partial differential equation (PDE) for the futures price

$$\begin{aligned} A'(\tau) + \sum_{i=1,2,3,v} B_i'(\tau) X_{i,t} + \sum_{i=1,2,3,v} B_i(\tau)(\mu_i^* + \kappa_i^* X_{i,t}) \\ + \frac{1}{2} \sum_{i,j=1,2,3,v} B_i(\tau)^2 V_t \sigma_{ij} + \sum_{i=1,2,3} \lambda_i \Big[\nu_i\big(B_i(\tau)\big) - 1\Big] = 0, \end{aligned} \tag{4.13}$$

Collecting terms in (4.13), the futures price must satisfy the following system of ordinary differential equations,

$$B_1'(\tau) = -\kappa_1^* B_1(\tau), \tag{4.14a}$$

$$B_2'(\tau) = -\kappa_2^* B_2(\tau), \tag{4.14b}$$

$$B_3'(\tau) = 0, \tag{4.14c}$$

$$\begin{aligned} B_v'(\tau) = &-B_v(\tau)\kappa_v^* + \frac{1}{2}\big(B_1(\tau)^2\sigma_1^2 + B_2(\tau)^2\sigma_2^2 + B_3(\tau)^2\sigma_3^2 \\ &+ B_v(\tau)^2\sigma_v^2 + 2B_1(\tau)B_2(\tau)\sigma_{12} + 2B_1(\tau)B_3(\tau)\sigma_{13} + 2B_1(\tau)B_v(\tau)\sigma_{1v} \\ &+ 2B_2(\tau)B_3(\tau)\sigma_{23} + 2B_2(\tau)B_v(\tau)\sigma_{2v} + 2B_3(\tau)B_v(\tau)\sigma_{3v}\big), \end{aligned} \tag{4.14d}$$

$$\begin{aligned} A'(\tau) = &-B_1(\tau)\alpha_1 - B_2(\tau)\alpha_2 + B_3(\tau)\mu_3^* + B_v(\tau)\kappa_v^*\theta_v^* + \lambda_1\big[\nu_1(B_1(\tau)) - 1 \\ &- B_1(\tau)(\nu_1(1) - 1)\big] + \lambda_2\big[\nu_2(B_2(\tau)) - 1 - B_2(\tau)(\nu_2(1) - 1)\big] \\ &+ \lambda_3\big[\nu_3(B_3(\tau)) - 1 - B_3(\tau)(\nu_3(1) - 1)\big], \end{aligned} \tag{4.14e}$$

with boundary conditions, $A(0) = 0$ and $B(0) = (1, 1, 1, 0)'$. Although the solutions

$B_1(\tau) = e^{-\kappa_1^*\tau}$, $B_2(\tau) = e^{-\kappa_2^*\tau}$, $B_3(\tau) = 1$ can be found easily, the ODE for $B_v(\tau)$ is of the Ricatti type. As the coefficients in the ODE for $B_v(\tau)$ are non-constant and depend on $\tau$, there is no known closed-form solution for $B_v(\tau)$ and subsequently $A(\tau)$. However, there are efficient methods for finding numerical solutions to the two ODE's such as Runge Kutta methods. This is the approach taken here for estimating the solutions to both futures and options prices.

For the nested Gaussian three-factor model and jump-diffusion model with no stochastic volatility, the futures price does have a closed form solution. This is shown in Appendix 4.A.

### 4.2.2 European Options on Futures

We now derive expressions for the price of European put and call options under the SVJ model. Consider an European call option on a futures contract where the futures contract expires at time $T_F$ and the option expires at time $T$, $T < T_F$. The dynamics of the futures price process under the risk-neutral measure can be derived using the generalized Itô's lemma as,

$$\begin{aligned}\frac{dF(t,T_F)}{F(t,T_F)} =& \sqrt{V_t}\big(B_1(\tau_F)\sigma_1 dZ_{1,t}^* + B_2(\tau_F)\sigma_2 dZ_{2,t}^* + B_3(\tau_F)\sigma_3 dZ_{3,t}^* + B_v(\tau_F)\sigma_v dZ_{v,t}^*\big) \\ &+ \Big(\exp\big\{e^{-\kappa_1^*\tau}J_{1,t}\big\} - 1\Big)dN_{1,t} - \lambda_1\Big(\frac{\omega_u}{\omega_u - e^{-\kappa_1^*\tau}} - 1\Big)dt \\ &+ \Big(\exp\big\{e^{-\kappa_2^*\tau}J_{2,t}\big\} - 1\Big)dN_{2,t} - \lambda_2\Big(\frac{\omega_d}{\omega_d + e^{-\kappa_2^*\tau}} - 1\Big)dt \\ &+ \Big(e^{J_{3,t}} - 1\Big)dN_{3,t} - \lambda_3\Big(e^{\sigma_j^2/2} - 1\Big)dt,\end{aligned} \tag{4.15}$$

where $\tau_F = T_F - t$. As it is easier to deal with the log-futures price when calculating option prices, we set $Y_t = \log F(t, T_F)$. Again, using the generalized Itô's lemma,

the dynamics of $Y_t$ is given by the following SDE,

$$\begin{aligned} dY_t = & \big(A'(\tau_F) + B'(\tau_F)X_t\big)dt + B(\tau_F)\cdot dX_{t_-} + (Y_t - Y_{t_-})dN_t \\ = & -\frac{1}{2}V_t\sigma_F(\tau_F)^2dt + \sqrt{V_t}\big(B_1(\tau_F)\sigma_1 dZ^*_{1,t} + B_2(\tau_F)\sigma_2 dZ^*_{2,t} + B_3(\tau_F)\sigma_3 dZ^*_{3,t} \\ & + B_v(\tau_F)\sigma_v dZ^*_{v,t}\big) + e^{-\kappa_1^*\tau_F}J_{1,t}dN_{1,t} - \lambda_1\left(\frac{\omega_u}{\omega_u - e^{-\kappa_1^*\tau_F}} - 1\right)dt \\ & + e^{-\kappa_2^*\tau_F}J_{2,t}dN_{2,t} - \lambda_2\left(\frac{\omega_d}{\omega_d + e^{-\kappa_2^*\tau_F}} - 1\right)dt + J_{3,t}dN_{3,t} - \lambda_3\left(e^{\sigma_j^2/2} - 1\right)dt, \end{aligned} \tag{4.16}$$

where

$$\begin{aligned} \sigma_F(\tau_F)^2 = & B_1(\tau_F)^2\sigma_1^2 + B_2(\tau_F)^2\sigma_2^2 + B_3(\tau_F)^2\sigma_3^2 + B_v(\tau_F)^2\sigma_v^2 \\ & + 2B_1(\tau_F)B_2(\tau_F)\sigma_{12} + 2B_1(\tau_F)B_3(\tau_F)\sigma_{13} + 2B_1(\tau_F)B_v(\tau_F)\sigma_{1v} \\ & + 2B_2(\tau_F)B_3(\tau_F)\sigma_{23} + 2B_2(\tau_F)B_v(\tau_F)\sigma_{2v} + 2B_3(\tau_F)B_v(\tau_F)\sigma_{3v}, \end{aligned} \tag{4.17}$$

Using the risk-neutral pricing framework, the price of the European futures call option at $t$ maturing at $T$ is given by the following expectation under the risk-neutral measure,

$$c = \mathbb{E}_{\mathbb{Q}}\big[e^{\int_0^T r_s ds}(F(T,T_F) - K)^+ | \mathcal{F}_t\big], \tag{4.18}$$

where $r_t$ denotes the instantaneous interest rate at time $t$. As it is assumed that interest rates are independent of the asset price under this modelling framework[33], the bond prices can be taken out of the expectation and we can represent the option price as,

$$\begin{aligned} c_t &= P(t,T)\mathbb{E}_{\mathbb{Q}}\big[(F(T,T_F) - K)\mathbf{1}_{F(T,T_F)>K} | \mathcal{F}_t\big] \\ &= P(t,T)\Big(\mathbb{E}_{\mathbb{Q}}\big[F(T,T_F)\mathbf{1}_{F(T,T_F)>K} | \mathcal{F}_t\big] - K\mathbb{E}_{\mathbb{Q}}\big[\mathbf{1}_{F(T,T_F)>K} | \mathcal{F}_t\big]\Big), \end{aligned} \tag{4.19}$$

[33]This is a reasonable assumption for commodity markets (including the crude oil market) since the volatility of interest rates has been found to be very small compared to the volatility of spot commodity prices by e.g. Schwartz (1997) and Casassus and Collin-Dufresne (2005).

where $P(t,T)$ is the value of a zero-coupon price maturing at time $T$. This expression can be evaluated by Fourier inversion of the characteristic function. Computing option prices by fourier inversion was introduced by Heston (1993) for stochastic volatility models and has been further expanded by Duffie et al. (2000) to accommodate general affine jump-diffusion models with jumps and stochastic volatility. A similar approach is used here given that the SVJ model is also an affine jump-diffusion model. This requires deriving the characteristic function for the futures price (or equivalently the log futures price) in the expectations above. The characteristic function for the log futures price maturing at time $T_F$ is defined as

$$\phi(\log F(t,T_F), V_t, t, T; u) = \mathbb{E}_{\mathbb{Q}}\big[e^{iu \log F(T,T_F)} | \mathcal{F}_t\big], \tag{4.20}$$

where $i = \sqrt{-1}$. From the transform analysis in Duffie et al. (2000), the characteristic function $\phi(\log F(t,T_F), V_t, t, T; u)$, for the log-futures price $Y_t$, has the following functional form

$$\phi(Y_t, V_t, t, T; u) = \exp\big\{M(\tau) + N(\tau)V_t + iuY_t\big\}, \tag{4.21}$$

where $\tau = T - t$ and with the boundary condition $\phi(Y_T, V_T, T, T; u) = e^{iu \log F(T,T_F)}$. Using a slight abuse of notation by letting $\phi_t = \phi(Y_t, V_t, t, T; u)$, the generalized Itô's lemma gives the dynamics of $\phi_t$ as

$$\begin{aligned}\frac{d\phi_t}{\phi_t} &= \big(-M'(\tau) - N'(\tau)V_t\big)dt + N(\tau)dV_t + iudY_{t-} + \frac{1}{2}\Big(N(\tau)^2 d\langle V\rangle_t \\ &\quad - u^2 d\langle Y^c\rangle_t + 2iuN(\tau)d\langle Y^c, V\rangle_t\Big) + (\phi_t - \phi_{t-})\end{aligned}$$

$$
\begin{aligned}
&= \Big[-M'(\tau) - N'(\tau)V_t + N(\tau)\kappa_v^*(\theta_v - V_t) + iu\sigma_{Fv}(\tau_F)N(\tau) + \frac{1}{2}\Big(N(\tau)^2\sigma_v^2 V_t \\
&\quad + iu(iu-1)\sigma_F(\tau_F)^2 V_t\Big)\Big]dt + \sqrt{V_t}\big(B_1(\tau_F)\sigma_1 dZ_{1,t}^* + B_2(\tau_F)\sigma_2 dZ_{2,t}^* \\
&\quad + (B_v(\tau_F) + N(\tau))\sigma_v dZ_{v,t}^*\big) + \Big(\exp\big\{iue^{-\kappa_1^*\tau_F}J_{1,t}\big\} - 1\Big)dN_{1,t} \\
&\quad + \Big(\exp\big\{iue^{-\kappa_2^*\tau_F}J_{2,t}\big\} - 1\Big)dN_{2,t} + \Big(\exp\big\{iuJ_{3,t}\big\} - 1\Big)dN_{3,t},
\end{aligned} \tag{4.22}
$$

where $Y_t^c$ denotes the continuous part of $Y$ and

$$
\sigma_{Fv}(\tau_F) = B_1(\tau_F)\sigma_{1v} + B_2(\tau_F)\sigma_{2v} + B_3(\tau_F)\sigma_{3v} + B_v(\tau_F)\sigma_v^2. \tag{4.23}
$$

Here, $\sigma_{Fv}(\tau_F)$ is the instantaneous covariance of the log-futures price with $V_t$ and $\sigma_F(\tau_F)^2$ defined by (4.17) is the instantaneous variance of the continuous part of the log-futures price. Using the fact that the characteristic function is a martingale under $\mathbb{Q}$ so that the drift in (4.22) is zero leads to the following equation for $M(\tau)$ and $N(\tau)$,

$$
\begin{aligned}
M'(\tau) + N'(\tau)V_t = N(\tau)\kappa_v^*(\theta_v^* - V_t) + iu\sigma_{Fv}(\tau_F)N(\tau) \\
+ \frac{1}{2}\Big(N(\tau)^2\sigma_v^2 V_t + iu(iu-1)\sigma_F(\tau_F)^2 V_t\Big).
\end{aligned} \tag{4.24}
$$

Collecting terms and noting that $\tau_F = \tau + T_f - T = \tau + \bar{\tau}$, it can be seen that $M(\tau)$ and $N(\tau)$ satisfy the following system of ODE's,

$$
\begin{aligned}
M'(\tau) &= N(\tau)\kappa_v^*\theta_v^* + \lambda_1\Big(\nu_1\big(iuB_1(\tau+\bar{\tau})\big) - 1 - iu\big(\nu_1(B_1(\tau+\bar{\tau}) - 1)\big)\Big) \\
&\quad + \lambda_2\Big(\nu_2\big(iuB_2(\tau+\bar{\tau})\big) - 1 - iu\big(\nu_2(B_2(\tau+\bar{\tau}) - 1)\big)\Big) \\
&\quad + \lambda_3\Big(\nu_3\big(iuB_3(\tau+\bar{\tau})\big) - 1 - iu\big(\nu_3(B_3(\tau+\bar{\tau}) - 1)\big)\Big),
\end{aligned} \tag{4.25a}
$$

$$
N'(\tau) = -N(\tau)\kappa_v^* + \frac{1}{2}N(\tau)^2\sigma_v^2 + iu\sigma_{Fv}(\tau+\bar{\tau})N(\tau) + \frac{1}{2}iu(iu-1)\sigma_F(\tau+\bar{\tau})^2, \tag{4.25b}
$$

with boundary conditions $M(0) = 0$ and $N(0) = 0$. This system of ODE's is again of the Riccati type. However, similar to the futures price, the coefficients in the Riccati ODE's partially depend on the maturity of the option, $\tau$, and so the solutions to both $M(\tau)$ and $N(\tau)$ do not exist. Numerical methods are therefore required to solve for both $M(\tau)$ and $N(\tau)$ to obtain the characteristic function. For the nested jump-diffusion model with no stochastic volatility, as with futures prices, the characteristic function has a closed-form solution. This is also demonstrated in Appendix 4.A.

European call and put options can then be priced using the Fourier inversion theorem considered by Heston (1993) and expanded by Bakshi and Madan (2000). We firstly introduce the following well-known result. For a random variable, $X$, with density function, $f(x)$ and characteristic function, $\phi(X;u)$ defined by

$$\phi(X;u) = \mathbb{E}\big[e^{iuX}\big] = \int_{-\infty}^{\infty} e^{iux} f(x)dx, \tag{4.26}$$

then

$$P(X > a) = \frac{1}{2} + \frac{1}{\pi}\int_0^{\infty} \mathrm{Re}\left[\frac{e^{-iua}\phi(X;u)}{iu}\right] du. \tag{4.27}$$

Using this result and following Bakshi and Madan (2000), the value of a call option is therefore given by,

$$c_t = P(t,T)(F(t,T_F)\Pi_1 - K\Pi_2), \tag{4.28}$$

where

$$\Pi_1 = \frac{1}{2} + \frac{1}{\pi}\int_0^{\infty} \mathrm{Re}\left[\frac{e^{-iu\log K - Y_t}\phi(Y_t, V_t, t, T; u-i)}{iu\phi(Y_t, V_t, t, T; -i)}\right] du,$$

$$\Pi_2 = \frac{1}{2} + \frac{1}{\pi}\int_0^\infty \text{Re}\left[\frac{e^{-iu\log K}\phi(Y_t, V_t, t, T; u)}{iu}\right] du.$$

Furthermore, if European put option prices are required then these can be easily obtained using put-call parity.

## 4.3 Estimation using Futures Data

In this section, we determine whether incorporating either jumps or volatility improves futures pricing performance in the crude oil futures market. Firstly, the estimation procedure for the models is introduced where due to the difficulty in estimating the stochastic volatility jump-diffusion model, only the SV and JD models are estimated. Next, the data used in the empirical study is described and the results of the estimation are presented where the impact of stochastic volatility or jumps on the model dynamics is analysed. This includes determining whether the extra dynamics improve futures pricing performance as well as whether they are able to better explain the distribution of crude oil prices. The last part in the section looks at the option pricing performance of each model when the option prices are computed using parameters estimated using futures prices only.

### 4.3.1 Estimation

Estimating stochastic volatility and jump-diffusion models in continuous-time is an ongoing area of research in financial modelling. Introducing these features causes the models to become nonlinear and/or non-Gaussian which complicates estimation. This is especially true for SV models when the volatility factor is governed

by a square-root diffusion process which is known to follow a non-central chi-square distribution. Additionally, for multi-factor models such as those used in this study, there are a large number of parameters, which motivates the need for computationally efficient estimation procedures. Due to this, many studies consider maximum likelihood estimation procedures which tend to be more computationally efficient than Bayesian methods such as MCMC whilst still giving reasonable results. A further complication with estimating continuous-time models is that data is observed discretely, which means that the model needs to be discretized to allow estimation.

The simplest and most popular method for discretizing continuous-time models is the Euler-Maruyama (EM) approach. Whilst the EM scheme has a lower order of convergence than other schemes, it is simple to implement and the bias tends to be small when observations are observed at a daily or higher frequency. A few studies have considered more accurate schemes for estimating continuous-time models including SV models and have compared them to the EM procedure. Lamoureux and Paseka (2009) derived a semi closed-form expression for the joint transition density of returns and the volatility process under the Heston (1993) model in terms of a single integral of Bessel functions. They compared their scheme to the EM scheme and found that although there was some bias when observations were infrequently sampled (> weekly), the discretization bias resulting from EM discretization was small when sampling at a daily or higher frequency. Aït-Sahalia (2002) derived an approximate analytical expression for the maximum likelihood estimator (MLE) for discretely sampled continuous-time models. Although his approach clearly outperformed the EM scheme for parameter estimation, the bias was not large[34]. Johannes et al. (2009) found that using EM discretization on stochastic volatility jump-diffusion models with infrequently sampled data (> weekly) results

[34] Aït-Sahalia (2002) found that his procedure was highly accurate to within a number of decimal places compared to a set of true ML parameter values. This was much better than estimation using EM, although EM was still accurate to within a couple of decimal places of the true ML parameter values.

in poor detection of jumps and daily or higher observations are required. They also found that the use of data augmentation[35] substantially reduces discretization bias on weekly or lower frequencies. However, for daily or higher frequencies, there is no economic difference in the filtering results when data augmentation is applied over no augmentation. Hence from these studies of both stochastic volatility and jump-diffusion models, the use of an EM discretization is sufficient for daily or higher frequencies. The size of each time-step for observations is therefore set to a daily frequency with the assumption that the size of each time-step is $\Delta t = 1/252$. This assumes a year consists of 252 trading days.

Using an EM discretization, the model is cast into a state space form. As mentioned in previous chapters, the state space model consists of a state transition equation which maps the state variables and errors to future states; and a measurement equation which relates the states to the observations. Although we present the state space form for the SVJ model here, only the nested models will be considered for estimation, namely the SV, JD and Gaussian models[36]. For the state transition equation, (4.3) can be discretized using the EM procedure as follows,

$$X_t = X_{t-1} + (\mu + \kappa X_{t-1})\Delta t + \sqrt{V_{t-1}\Delta t}\epsilon_t + J_t \Delta N_t, \quad \epsilon_t \sim N(0, \Sigma\Sigma'), \tag{4.30}$$

where $\mu$, $\kappa$, $\Sigma$ and $J_t$ come from (4.3), $\epsilon_t$ is a $\mathbb{R}^4$ conditionally Gaussian random variable with mean 0 and variance $\Sigma\Sigma'$, $N_t$ is a $\mathbb{R}^4$ Poisson process and $\Delta N_t = N_t - N_{t-1}$.

For the measurement equation, as more maturities are observed than state variables to be estimated, the futures prices are assumed to be observed with error. To help detect any model misspecification, it is assumed that the errors

[35] Johannes et al. (2009) describe data augmentation as simulating prices between observations during filtering to reduce discretization bias.

[36] The SVJ model is estimated later using options data.

are modelled as AR(1) Gaussian errors similar to the previous chapter. Letting $Y_t = (\log F(t, t+T_1), \ldots, \log F(t, t+T_N))'$, the observations and states are related through the following expressions,

$$\begin{aligned} Y_t &= A_t + B_t X_t + e_t, \\ e_t &= \phi e_{t-1} + \nu_t, \end{aligned} \tag{4.31}$$

where $A_t = (A(T_1 - t), \ldots, A(T_N - t))'$, $B = (B(T_1 - t), \ldots, B(T_N - t))'$, $\nu_t$ are $\mathbb{R}^N$ serially uncorrelated $\mathrm{N}(0, H_t)$. For simplicity it is assumed that $\phi$ is the same for all contracts and that the measurement errors are uncorrelated such that $H_t = \mathrm{diag}(s_1^2, \ldots, s_N^2)$. To accommodate the AR(1) measurement errors in the state space model, they must be placed in the state equation. This requires a slight modification to the state space model with the state variable defined as $x_t = (X_{1,t}, X_{2,t}, X_{3,t}, V_t, \epsilon_t)$' and the state transition equation as

$$x_t = T_t x_{t-1} + c_t + \eta_t, \tag{4.32}$$

where $T_t = \begin{pmatrix} 1 - \kappa\Delta t & 0 \\ 0 & \phi \end{pmatrix}$, $T_t \in \mathbb{R}^{(4+N)\times(4+N)}$, $c_t = (\mu\Delta t, 0)'$, $c_t \in \mathbb{R}^{4+N}$ and $\eta_t = (\epsilon_t, \nu_t)'$, $\eta_t \in \mathbb{R}^{(4+N)}$.

Due to the nature of the nonlinearity introduced by the SV model, it is difficult to implement more efficient maximum likelihood procedures such as the MCL procedure[37] or the smooth particle filter due to the multi-dimensional model. Also, given the number of parameters to be estimated and the desire for computational efficiency, the choice was made to use a quasi-maximum likelihood (QML) procedure to estimate the stochastic volatility model. Under this procedure, although the distribution of each state variable depends on the volatility factor $V_t$, it is assumed

[37] However, the MCL procedure is used to estimate the jump-diffusion model with no stochastic volatility.

that the state variables are conditionally Gaussian. This assumption means that the likelihood function can be estimated using the Kalman filter recursions which were outlined in Chapter 1, Section 1.2.

Whilst this approach is computationally efficient, the resulting maximum likelihood estimator is the minimum linear mean-squared estimator (rather than the minimum mean-square estimator) and is inconsistent. A modification is also needed to be applied to the Kalman filter recursions in order to incorporate the stochastic volatility factor. As the covariance matrix of the state errors are assumed to depend on $V_{t-1}$, the covariance matrix must be updated during each time step. This is essentially the same as applying the extended Kalman filter (EKF) although in this case the approach is specific to this model. For brevity, we do not include the EKF recursions for general nonlinear state space models here, although it requires only minor modifications to the standard linear Kalman filter by linearising the model through a first-order Taylor expansion of the state space model. Also, an issue with making a Gaussian assumption for $V_t$ is that the Kalman filter equations can result in a negative value for $V_t$ despite the parameter restrictions in place. In cases where $V_t$ becomes negative, $V_t$ is set to $-V_t$ to ensure that it is both positive and non-zero[38].

Although the use of QML results in biased estimates, studies in the term structure literature tend to suggest that the bias is small when estimating the square-root diffusion process using QML (e.g. Duan and Simonato (1999) and Chen and Scott (2003)). Also, when incorporating the stochastic volatility factor in the state equation, the Kalman filter has been found to perform reasonably well in practice for Heston-type volatility models (e.g. Bibkov and Chernov (2009) and Trolle and Schwartz (2009) among others). Based on these studies, there is no reason to believe that a large bias in the estimator exists as a result of using QML.

[38]This is only an issue when $V_t$ approaches 0. For the parameters estimated using the crude oil futures data, $V_t$ was never close to zero or became negative and so this was not an issue.

As a check, a small finite-sample study was conducted on the QML procedure based on an EM discretization of the model described by (4.1) and the corresponding futures price given by (4.9) and (4.14a). The study consists of simulations of the state variables and futures prices with maturities of up to 2 years and uncorrelated normally distributed measurement errors. The length of each time series was set to 2500 daily observations and a cross-section of 7 futures contracts with maturities of 1, 3, 6, 9, 12, 18 and 24 months. A total of 400 simulations were conducted using parameters set close to the parameters estimated from the real data set. The model parameters were then estimated using the QML procedure for each simulated data set using a range of starting values for the parameters.

**Table 4.1**

Finite Sample performance of the QML estimator

This table reports the estimation results for simulations of the model given by (4.30) to produce a cross-section of futures prices with measurement errors given by (4.31). Each simulation consists of 2500 daily observations and 400 simulations were conducted. The parameters used to generate the simulations are listed in the rows labelled "True". The table also lists the mean parameter estimates based on QML estimation of each simulated sample path and the root mean squared error (RMSE) of the estimated parameters from the true parameters.

| Parameter | True | Mean | RMSE | Parameter | True | Mean | RMSE |
|---|---|---|---|---|---|---|---|
| $\kappa_1^*$ | 2.0000 | 1.9980 | 0.0263 | $\rho_{12}$ | −0.1000 | −0.1001 | 0.0169 |
| $\alpha_1$ | 0.1000 | 0.0898 | 0.0322 | $\rho_{13}$ | 0.0000 | 0.0021 | 0.0146 |
| $\sigma_1$ | 0.2500 | 0.2505 | 0.0101 | $\rho_{1v}$ | 0.2000 | 0.1995 | 0.0132 |
| $\kappa_2^*$ | 1.0000 | 1.0009 | 0.0159 | $\rho_{23}$ | 0.0000 | −0.0010 | 0.0133 |
| $\alpha_2$ | 0.0500 | 0.0404 | 0.0286 | $\rho_{2v}$ | 0.1000 | 0.0985 | 0.0161 |
| $\sigma_2$ | 0.2000 | 0.2000 | 0.0091 | $\rho_{3v}$ | 0.1500 | 0.1472 | 0.0116 |
| $\mu_3$ | 0.0500 | 0.0504 | 0.0163 | $s_1$ | 0.0010 | 0.0010 | 0.0001 |
| $\mu_3^*$ | −0.0500 | −0.0438 | 0.0139 | $s_2$ | 0.0010 | 0.0010 | 0.0000 |
| $\sigma_3$ | 0.3000 | 0.2892 | 0.0237 | $s_3$ | 0.0010 | 0.0010 | 0.0000 |
| $\kappa_v^*$ | 3.0000 | 2.9914 | 0.0998 | $s_4$ | 0.0010 | 0.0010 | 0.0000 |
| $\theta_v^*$ | 1.0000 | 1.0027 | 0.0312 | $s_5$ | 0.0010 | 0.0010 | 0.0000 |
| $\theta_v$ | 1.0000 | 1.0001 | 0.0330 | $s_6$ | 0.0010 | 0.0010 | 0.0000 |
| $\sigma_v$ | 0.5000 | 0.4819 | 0.0431 | $s_7$ | 0.0010 | 0.0010 | 0.0000 |

The results of the finite-sample performance are reported in Table 4.1. Overall the

bias was found to be quite small for most of the parameter estimates. The means for most parameters are close to the true parameter values except for some of the risk premium parameters. The RMSE of the parameters are also small relative to the values of the parameters. Hence, these results suggest that it is reasonable to use the QML procedure for parameter estimation of the SV model.

Additionally, for the jump-diffusion model with no stochastic volatility, estimating the model via the MCL procedure is possible if it is assumed that the state variables are uncorrelated (i.e. $\rho_{ij} = 0$, $i \neq j$). The approach was used in Dempster et al. (2008) and the same approach is used here. Details of the implementation for the jump-diffusion model can be found in Appendix 4.B.

As we have introduced the jump-diffusion and stochastic volatility model as an extension to the long/short three factor model, we compare the three specifications, the Gaussian, JD and SV models. To estimate the parameters, the likelihood function is maximized in the same manner as before using the Kalman filter recursions outlined in Chapter 1 and a BFGS optimization procedure provided in Matlab. The standard errors are computed using (2.11) in Chapter 2 which gives a consistent estimate of the standard errors for the QML procedure. Also, to compute the futures prices, numerical solutions to the ODE's in equation (4.14a) are needed. These are computed using the *ode45* function in Matlab which implements a Runge-Kutta (4,5) method for solving the ODE's. For the SV model, there are 22 parameters related to state variables to be estimated and for the jump-diffusion model with no stochastic volatility and uncorrelated state disturbances, there are 16 parameters. Also, as the jump-diffusion model does not include correlations between the three factors, only the Gaussian and SV model are nested specifications. However, comparisons can be made between all three models through the model pricing errors and the Akaike and Bayesian (Schwartz) information criterion (or AIC and BIC)

values.

### 4.3.2 Data

The data used for estimation consists of daily observations of the West Texas Intermediate (WTI) crude oil futures traded on the New York Mercantile Exchange (NYMEX). The data spans the period from January 2000 - March 2009 giving a total of 2200 observation dates. The contract maturities chosen include the following: 1, 3, 6, 9, 12, 18 and 24 month contracts giving a total of 7 different maturities. Contracts with longer maturites were also available, however only contracts with up to two years to maturity were trading on a daily basis with sufficient volume[39]. Also, we only include contracts with at least 10 trading days before the last day of trading in each contract for the same reasons as in the previous chapter. This is because during the final week of trading in futures contracts, open interest rapidly declines. With daily data, the microstructure effects of this is likely to be quite high and to mitigate these effects, these observations are not included for analysis. As the data consists of observations during the "global financial crisis" (GFC), a sub-sample prior to the peak of the crisis in September 2008 is also considered.

We provide some descriptive statistics of the annualized daily futures returns in Table 4.2. As expected for futures returns, the mean of the returns for all contracts are close to zero, although the standard errors are quite high indicating that futures returns are highly variable. The "Samuelson effect" is evident in daily futures returns as variance decreases with increasing maturity. The returns also exhibit negative skewness and leptokurtosis which suggests that they are non-normal and

[39] Although all contract settlement prices were adjusted daily, these maturities were found to be traded frequently enough to be used in the analysis for daily observations.

have fat tails. This is mostly indicated by the kurtosis of returns, which is quite high for all contracts. This gives some indication that incorporating jumps or stochastic volatility may improve modelling performance.

**Table 4.2**
Descriptive Statistics for Futures Returns, 2000-2009

This table provides descriptive statistics for the annualized daily continuously compounded returns (in percentage points) of crude oil futures contracts traded on NYMEX during the period Jan 2000 - Mar 2009.

| Contract | Mean Maturity (in months) | Mean (SE) | Variance (SE) | Skewness (SE) | Kurtosis (SE) |
|---|---|---|---|---|---|
| 1 | 0.84 | 0.52 | 40.40 | -0.21 | 6.43 |
| | | (0.84) | (48.14) | (0.05) | (0.10) |
| 2 | 2.85 | 0.59 | 35.64 | -0.26 | 5.59 |
| | | (0.74) | (37.47) | (0.05) | (0.10) |
| 3 | 5.87 | 0.69 | 31.74 | -0.27 | 5.94 |
| | | (0.66) | (29.72) | (0.05) | (0.10) |
| 4 | 8.89 | 0.75 | 29.50 | -0.22 | 5.86 |
| | | (0.62) | (25.68) | (0.05) | (0.10) |
| 5 | 11.90 | 0.80 | 27.94 | -0.21 | 5.86 |
| | | (0.58) | (23.02) | (0.05) | (0.10) |
| 6 | 17.94 | 0.87 | 26.05 | -0.21 | 5.83 |
| | | (0.54) | (20.02) | (0.05) | (0.10) |
| 7 | 23.99 | 0.91 | 25.75 | -0.17 | 6.91 |
| | | (0.54) | (19.56) | (0.05) | (0.10) |

### 4.3.3 Empirical Results

#### Parameter Estimates

Table 4.3 report the parameter estimates for the Gaussian, JD and SV models for the full-sample and the sub-sample period from 4 January 2000 - 29 August 2008. We will mainly concentrate on the spot price parameters for the full-sample but will briefly comment on whether the GFC has impacted the parameters much. Overall, most of the parameter estimates are significant for all of the models apart from

some of the risk premium terms. The risk-neutral mean-reversion parameters for all models seem to be fairly similar and as the futures term structure is mainly driven by these parameters, it suggests that the introduction of non-Gaussian dynamics does not affect the futures term structure much.

If we consider the risk-premium parameters, although the significance can be established for some of the constant MPR parameters, most of the time-varying MPR parameters are insignificant. A likely explanation for this is that using daily observations introduces additional noise in the spot price dynamics which makes it difficult to accurately estimate the risk-premium parameters. Although $\beta_1$ and $\beta_2$ were very different between each of the models, the insignificance of most of these parameters means meaningful comparisons cannot be made. However, the results do give some indication that the levels of mean-reversion differ under each of the model specifications. The most notable differences between the models is the real-world mean-reversion of $X_{1,t}$ as $\beta_1$ is significant for the JD and SV models with a value of 4.14 and -1.84 respectively. This does suggest that the behaviour of spot prices differs for each model under the real-world measure and that they imply different dynamics of risk premiums. Whilst this does pose some interesting questions with regard to the impact of jump risk and volatility risk on risk premiums, this is not the focus of this study and is left for future research consideration.

Turning to the jump-diffusion model, the parameters for all of the jumps are significant indicating that jumps are exhibited by spot prices. For the full sample, the frequency of short-term up jumps is around 8.6 per year, 18 per year for short-term down jumps and 48.9 per year for the long-term jumps. However, although there are a fairly large number of jumps exhibited, these jumps tend to be quite small, the mean jump sizes are 2.8% for $J_1$, 2% for $J_2$, and 2.4% for $J_3$. Although the down jump size are smaller on average than the up jumps, the down jumps occur

on average 2 times as often. This is consistent with the skewness exhibited by the futures returns. Also, the jump parameters for the sub-sample prior to September 2008 were not substantially affected apart from the long-term jumps. This suggests that the GFC caused an increase in large parallel shifts of the futures curve.

Examining the stochastic volatility factor, all of the volatility parameters are significant except for the long-term mean under the real-world measure, $\theta_v$. For the volatility parameter values under the risk-neutral measure, volatility exhibits quite strong mean-reversion with $\kappa_v^* = 6.05$. However, under the real-world measure, the mean-reversion is much lower with a value of $\kappa_v = \kappa_v^* - \beta_v = 0.57$. This suggests that volatility is quite persistent with low mean-reversion. An implication of the high value of $\beta_v$ is that there are significant volatility risk premiums present. This is also supported by the value of $\alpha_v = \kappa_v\theta_v - \kappa_v^*\theta_v^* = -0.40$ which suggests negative volatility risk premiums are present. The volatility of volatility parameter, $\sigma_v$ is quite high with a value of 1.166 which suggests that volatility does tend to vary substantially over time. However, this value does appear to be inflated as a result of the GFC as the value of $\sigma_v$ estimated on the sub-sample period is only 0.72.

It is also worth commenting on the correlation parameters for the SV model. The parameter estimation results show that there are significant correlations between the spot price factors and the volatility factor. It has been suggested in the literature (e.g. Geman (2005), pg. 105) that contrary to the behaviour of equities, an inverse "leverage effect" is present in commodity markets where volatility is positively related to the spot price. This comes from the view that higher commodity prices negatively impact the world economy. From a simple calculation, the instantaneous correlation coefficient of the spot price with volatility can be calculated as 0.06. This suggests that there is a very weak inverse leverage effect, although it is suspected that the behaviour of crude oil prices during the peak of the GFC may

**Table 4.3**

Parameter Estimates

This table reports the parameter estimates for the short/long term Gaussian, jump-diffusion and stochastic volatility model obtained by maximum likelihood. The parameters are defined by the SDE's in (4.1) and (4.6) cast into the state space model described by (4.30) - (4.32). The full data sample used consists of daily futures prices for the period 4 January 2000 - 24 March 2009 with maturities of approximately 1, 3, 6, 9, 12, 18, and 24 months. The sub-sample consists of observations for the period 4 January 2000 - 29 August 2008. S.E. denotes the standard errors of the parameter estimates.

| | Jan 2000 - Mar 2009 | | | | | | Jan 2000 - Aug 2008 | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Gaussian | | JD | | SV | | Gaussian | | JD | | SV | |
| Parameter | Estimate | S.E. | Estimate | S.E. | Estimate | S.E. | Estimate | S.E. | Estimate | S.E. | Estimate | S.E. |
| $\kappa_1^*$ | 8.5861 | (0.1574) | 8.6815 | (0.1565) | 8.2713 | (0.0186) | 7.6166 | (0.1995) | 7.8287 | (0.1846) | 7.6660 | (0.0482) |
| $\alpha_1$ | −0.2487 | (0.0851) | −0.6660 | (0.1527) | −0.5087 | (0.0286) | −0.1891 | (0.0679) | −0.4451 | (0.0692) | −0.3680 | (0.1600) |
| $\beta_1$ | 1.3482 | (1.2898) | 4.1360 | (1.0373) | −1.8356 | (0.7113) | −0.4120 | (1.3353) | −0.5150 | (1.2858) | −3.6847 | (0.7324) |
| $\sigma_1$ | 0.2138 | (0.0041) | 0.1821 | (0.0046) | 0.2495 | (0.0020) | 0.1747 | (0.0040) | 0.1527 | (0.0048) | 0.2723 | (0.0014) |
| $\kappa_2^*$ | 1.1246 | (0.0216) | 1.0985 | (0.0190) | 1.1018 | (0.0346) | 1.1899 | (0.0263) | 1.1518 | (0.0208) | 1.1920 | (0.0241) |
| $\alpha_2$ | 0.0593 | (0.0927) | 0.8040 | (0.6336) | −0.5424 | (0.3144) | 0.0855 | (0.0785) | 0.4442 | (0.1714) | 0.1345 | (0.1188) |
| $\beta_2$ | 0.0571 | (0.4598) | 0.4789 | (0.5093) | 0.3683 | (0.4245) | −0.1338 | (0.7041) | 0.2807 | (0.3528) | 0.0806 | (0.2024) |
| $\sigma_2$ | 0.2579 | (0.0039) | 0.2271 | (0.0053) | 0.3156 | (0.0065) | 0.2777 | (0.0039) | 0.2114 | (0.0044) | 0.3680 | (0.0028) |
| $\mu_3$ | 0.1377 | (0.0814) | 0.2126 | (0.0729) | −0.0262 | (0.0919) | 0.2203 | (0.0751) | 0.2499 | (0.0721) | 0.0260 | (0.0761) |
| $\mu_3^*$ | −0.0472 | (0.0035) | −0.0361 | (0.0035) | −0.0490 | (0.0022) | −0.0518 | (0.0030) | −0.0459 | (0.0031) | −0.0465 | (0.0595) |
| $\sigma_3$ | 0.2457 | (0.0037) | 0.1862 | (0.0061) | 0.3167 | (0.0044) | 0.2197 | (0.0036) | 0.1909 | (0.0058) | 0.3887 | (0.0070) |
| $\kappa_v^*$ | – | – | – | – | 6.0460 | (0.0011) | – | – | – | – | 5.9191 | (0.0032) |
| $\theta_v^*$ | – | – | – | – | 0.1782 | (0.0926) | – | – | – | – | 0.3033 | (0.9050) |
| $\theta_v$ | – | – | – | – | 1.1839 | (0.0002) | – | – | – | – | 0.4017 | (0.0024) |
| $\beta_v$ | – | – | – | – | 5.4719 | (0.0006) | – | – | – | – | 5.2681 | (0.0088) |
| $\sigma_v$ | – | – | – | – | 1.1659 | (0.0002) | – | – | – | – | 0.7231 | (0.0030) |
| $\rho_{12}$ | 0.0417 | (0.0248) | – | – | 0.0255 | (0.0193) | 0.0045 | (0.0274) | – | – | −0.0162 | (0.0012) |
| $\rho_{13}$ | 0.0301 | (0.0220) | – | – | 0.0722 | (0.0218) | 0.1004 | (0.0232) | – | – | 0.0393 | (0.0020) |
| $\rho_{1v}$ | – | – | – | – | 0.7234 | (0.0565) | – | – | – | – | 0.6951 | (0.0246) |
| $\rho_{23}$ | 0.2338 | (0.0221) | – | – | 0.2464 | (0.0066) | 0.1749 | (0.0235) | – | – | 0.2605 | (0.0086) |
| $\rho_{2v}$ | – | – | – | – | −0.2664 | (0.0489) | – | – | – | – | 0.0510 | (0.0044) |
| $\rho_{3v}$ | – | – | – | – | −0.1502 | (0.0149) | – | – | – | – | −0.1456 | (0.0072) |
| $\lambda_1$ | – | – | 8.5832 | (2.3369) | – | – | – | – | 12.0673 | (2.3631) | – | – |
| $\omega_u$ | – | – | 35.4172 | (6.8063) | – | – | – | – | 53.0250 | (4.4989) | – | – |
| $\lambda_2$ | – | – | 17.9527 | (5.1437) | – | – | – | – | 11.8685 | (4.2939) | – | – |
| $\omega_d$ | – | – | 45.0809 | (6.5311) | – | – | – | – | 40.9106 | (13.7767) | – | – |
| $\lambda_3$ | – | – | 48.8705 | (8.6550) | – | – | – | – | 31.3209 | (8.3046) | – | – |
| $\sigma_j$ | – | – | 0.0233 | (0.0018) | – | – | – | – | 0.0193 | (0.0021) | – | – |
| $\phi$ | 0.9793 | (0.0022) | 0.9794 | (0.0025) | 0.9652 | (0.0071) | 0.9715 | (0.0026) | 0.9726 | (0.0029) | 0.9706 | (0.0103) |
| $s_1$ | 0.0009 | (0.0003) | 0.0008 | (0.0002) | 0.0009 | (0.0003) | 0.0011 | (0.0002) | 0.0010 | (0.0002) | 0.0010 | (0.0001) |
| $s_2$ | 0.0015 | (0.0001) | 0.0015 | (0.0000) | 0.0015 | (0.0001) | 0.0015 | (0.0000) | 0.0015 | (0.0000) | 0.0015 | (0.0000) |
| $s_3$ | 0.0006 | (0.0000) | 0.0006 | (0.0000) | 0.0006 | (0.0001) | 0.0006 | (0.0000) | 0.0006 | (0.0000) | 0.0006 | (0.0000) |
| $s_4$ | 0.0005 | (0.0000) | 0.0005 | (0.0000) | 0.0005 | (0.0000) | 0.0004 | (0.0000) | 0.0004 | (0.0000) | 0.0004 | (0.0000) |
| $s_5$ | 0.0007 | (0.0000) | 0.0007 | (0.0000) | 0.0007 | (0.0001) | 0.0007 | (0.0000) | 0.0007 | (0.0000) | 0.0007 | (0.0000) |
| $s_6$ | 0.0018 | (0.0000) | 0.0018 | (0.0000) | 0.0018 | (0.0001) | 0.0018 | (0.0000) | 0.0018 | (0.0000) | 0.0018 | (0.0000) |
| $s_7$ | 0.0046 | (0.0001) | 0.0046 | (0.0001) | 0.0046 | (0.0001) | 0.0046 | (0.0001) | 0.0046 | (0.0001) | 0.0046 | (0.0001) |
| Loglik | 69016.96 | | 69306.44 | | 69604.05 | | 65568.45 | | 65752.73 | | 65864.57 | |

have affected this result. During the peak of the GFC, commodity prices (along with all other assets) worldwide collapsed when volatility was extremely high which implies a negative correlation between volatility and prices. However, the instantaneous correlation of volatility and spot prices for the sub-sample period of January 2000 - August 2008 is only slightly higher with a value of 0.12. As the correlation is quite low, this does not provide strong evidence for the existence of an inverse leverage effect in crude oil prices.

It is possible to see how well the SV model is able to capture the volatility of futures prices. A plot of the estimated volatility for the nearby futures contract can be found in Figure 4.1. In the figure, historical volatility is computed using the discrete-time SV model considered in Chapter 2[40]. The volatility of the nearby futures contract under this chapter's SV model is given by (4.17). The plot shows that the estimated volatility tracks historical volatility reasonably well, although the fit appears to be much better during the second half of the sample from about 2005 onwards. The sample correlation between historical volatility and the estimated volatility is about 0.73 which supports what can be observed in the plot. This demonstrates that the model dynamics may be sufficient to capture volatility in the crude oil futures market although it is unclear whether a more accurate estimation procedure than QML or a better specified model will improve the estimates.

These results are somewhat contradictory to Hughen (2010), who found little correlation between the volatility of the nearby crude oil futures contract and the estimated volatility of the SV model in his paper. He used this result to provide

[40]The most appropriate measure for daily volatility is arguably a realized volatility estimate. However, as intra-daily data was not available a compromise is the volatility estimated under the discrete-time SV model of Chapter 2. Even though we do not know what the true model of volatility is, from the analysis in Chapter 2, the discrete-time model was found to estimate crude oil volatility reasonably well. Comparisons (not reported) with other volatility measures such as GARCH or rolling window measures gave similar volatility estimates. Hence, it seems reasonable to assume that historical volatility can be represented by the volatility estimated under the discrete-time SV model.

evidence of "unspanned stochastic volatility" (USV), a term first introduced in the term strucutre literature. Unspanned stochastic volatility can be defined as markets which can not be completed using futures (bonds) only so that derivatives can not be perfectly replicated using futures (bonds) in commodity (fixed income) markets. However, poor correlation between the estimated volatility and historical volatility does not necessarily imply USV in this case. Firstly, Hughen (2010) only used two spot price factors to price futures contracts. As demonstrated in the previous chapter, two spot price factors are not sufficient for modelling the futures term structure. This means that the volatility factor in Hughen (2010) was used in a dual role as a substitute for an additional spot price factor for futures pricing and to model volatility. This is in contrast to this study where three spot price factors are used to capture the futures term structure, leaving the volatility factor to mainly capture volatility. In fact, when Hughen (2010) used a USV version of the model where the volatiltiy factor has no impact on futures pricing, he found a similar correlation coefficient (0.74 for historical volatility) to what was observed here. However, this does not preclude that USV exists in the crude oil market, although we do not deal with USV directly in this study. This issue is discussed more in the next section in relation to option pricing.

### Model Performance and Futures Pricing

We firstly compare the futures pricing performance based on the estimation results above. Although the SV and Gaussian models are nested, the JD model is not nested with the other models, meaning direct comparison of likelihood values is limited. However, we can compare the Akaike and Bayesian (Schwartz) Information Criterion or AIC and BIC values for all three models. These values are reported in Table 4.4 where it can be seen that the SV model has the lowest values for both metrics by

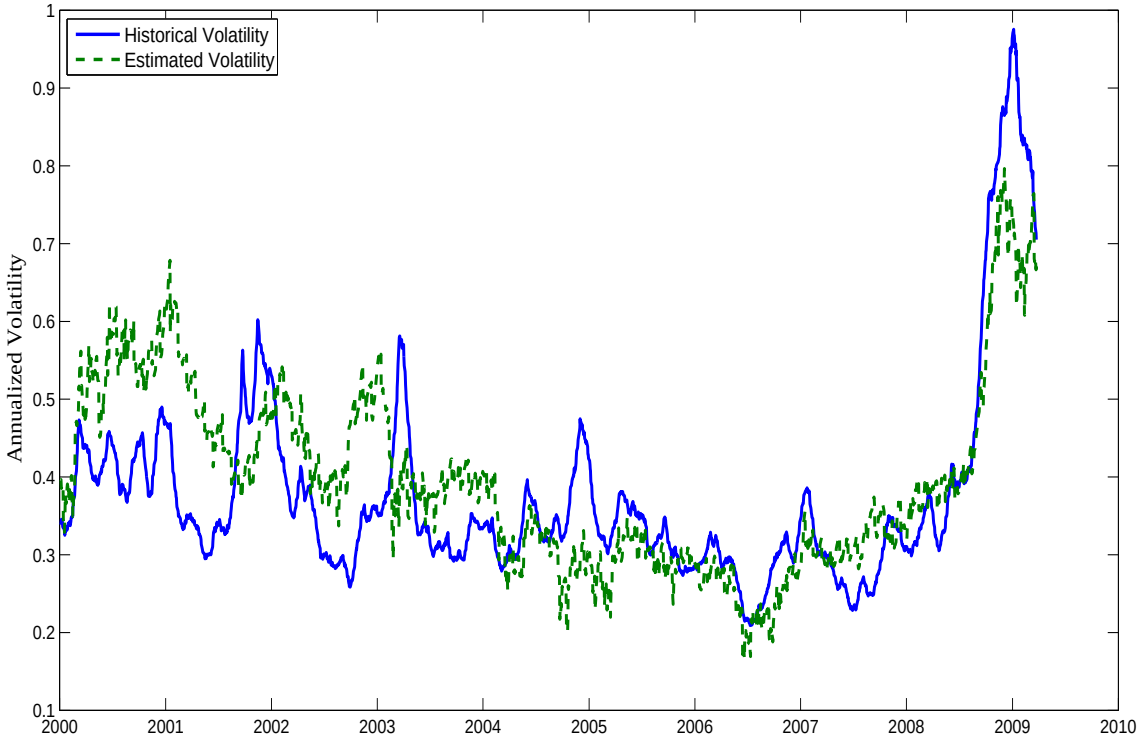



**Figure 4.1**
Plot of historical volatility and estimated volatility on the nearby futures contract, Jan 2000 - March 2009

a substantial margin, followed by the JD model. This suggests that incorporating either stochastic volatility or jumps in the dynamics improves the goodness of fit. A likelihood ratio (LR) test between the SV model and the Gaussian model is also possible[41]. The LR test is conducted for the restricted Gaussian model and the unrestricted SV model where the restricted model has 8 parameter restrictions. The LR test statistic given the log-likelihoods in Table 4.3 is 1174.18 which is much larger than the $\chi^2_{0.05,8}$ value of 15.51. Hence the hypothesis that there is no stochastic volatility is strongly rejected.

To compare the futures pricing performance of each model, the in-sample estimates of the futures prices are obtained from the filtered values of the state variables for

[41] Whilst the JD model could possibly be compared using the Vuong (1989) LR test for non-nested models, it should be noted that the MCL procedure does not provide the marginal likelihoods. This means it is not possible to conduct the LR test. Thus we only consider the informal comparisons for the JD model.

each model. The daily root mean-squared errors (RMSE)[42] of each model are then computed from the cross-section of log-futures prices. The values in Table 4.4 report the mean of the daily RMSE values. Interestingly, the Gaussian model is able to provide the best fit to the futures prices, although there is not a large difference between each of the three specifications. This result is not completely unexpected given that the the extra dynamics in the jump-diffusion and SV models have little effect on the first moment of the spot price. Hence, futures prices will tend to be similar despite the addition of non-Gaussian features. However, it is surprising that the Gaussian model performs the best. This indicates that whilst the underlying spot price dynamics are better explained by the non-Gaussian elements, this does not translate into better futures pricing performance.

It should also be noted that $\phi$ is quite high and similar for each model. All three models have an estimate around 0.97 indicating that the models are misspecified. Whilst this is of some concern, it should be noted that the high correlation in measurement errors can partly be explained by the noise in daily futures prices. This tends to be mitigated over weekly or monthly observations[43]. Also, given that the daily RMSE is less than 1% for all models, this does not seem to be a major issue as the models are still able to fit the futures curve well.

In summary, the improved performance of the two extended models, as shown by the AIC and BIC values and the LR test, is most likely due to their being able to better explain the distributional properties of the underlying state variables. Hence, although futures prices fit better under the Gaussian model, the distributional properties of the state variables implies that a Gaussian model is insufficient for modelling spot prices. This may also be an issue in relation to pricing more

[42]The daily RMSE is calculated from the cross-section of observed and estimated futures prices each day.

[43]This is the case in the previous chapter where the three-factor Gaussian model exhibits much lower values of $\phi$ when weekly observations were used.

complex derivatives such as options where the Gaussian model is likely to perform poorly compared to the other models. This leads to the next section where we compare the option pricing performance of the models.

**Table 4.4**
Model Errors on Futures Contracts

This table reports the AIC and BIC values for each model specification for the full sample and the sub-sample period. RMSE denotes the mean of the daily root-mean squared error of the difference between the observed log futures prices and the estimated log futures prices. The values of the errors are reported in percentage points.

| | Jan 2000 - Mar 2009 | | | Jan 2000 - Aug 2008 | | |
|---|---|---|---|---|---|---|
| Model | AIC | BIC | RMSE | AIC | BIC | RMSE |
| Gaussian | −135319.19 | −135150.10 | 0.67 | −129561.65 | −129393.95 | 0.58 |
| JD | −135761.08 | −135568.92 | 0.79 | −129849.76 | −129659.18 | 0.70 |
| SV | −136159.01 | −135920.74 | 0.75 | −130131.55 | −129895.24 | 0.68 |

### Option Pricing

The options data used in the analysis consists of daily settlement prices of call and put options on WTI crude oil futures contracts traded on NYMEX and accessed through Datastream. However, only data from 30 August 2006 was available on Datastream. Also, on inspection of the data obtained, it was found that there was a lot of missing data for contracts maturing in September 2008 - December 2008[44]. As a result of this, only data from 30 August 2006 - 1 August 2008 is considered. This means that the options data set does not span the the same length as the futures data, although the sample lies within the futures data sample. Also, since there are a large number of option prices for a range of strike prices available, only a subset of the options data is used. The options are chosen at each date according to their maturity and moneyness, where moneyness is defined as the strike price divided by the futures price. This means for a given strike price $K$ and futures price $F_t$, call

[44]Most put option contracts were missing for September - December options and all call option data was missing for these contracts.

options are at-the-money (ATM) when $K = F_t$, in-the-money (ITM) when $K < F_t$ and out-of-the-money (OTM) when $K > F_t$ and vice-versa for put options. The moneyness categories chosen in this empirical study are 0.8, 0.85, 0.9, 0.95, 1, 1.05, 1.1, 1.15 and 1.2. The options closest to each moneyness category at each date and maturity are chosen to ensure consistency in the comparison over time. As trading in the options ends three business days before the termination of trading in the underlying futures contract, the maturity of the options is assumed to occur on the date when trading ceases. Also, to remain consistent with the futures data set, only options with at least 10 trading days to maturity are considered.

When comparing option pricing performance, the convention in the literature is to compare implied volatilities rather than prices as they tend to be more stable across maturities and strikes. Given that the options are on futures, the Black (1976) implied volatilities are to gauge option pricing performance. Also, the options traded are of the American type which means there is an early exercise premium that needs to be accounted for[45]. In order to estimate the early exercise premium, the implied volatilities are obtained by inverting the Barone-Adesi and Whaley (1987) American option pricing formula which uses a quadratic analytic approximation[46]. The equivalent European option prices were then obtained using the Black (1976) futures option formula with the estimated implied volatilities. Although other methods for computing American option prices are available such as binomial/trinomial trees, the Barone-Adesi and Whaley (1987) approach is chosen as it provides comparable accuracy to tree methods, is simple to implement and computationally more efficient. Similar approaches are taken by Richter and Sørensen (2002) and Trolle and Schwartz (2009) using commodity options and Broadie et al. (2007) using equity options to estimate European futures options prices when only American option

[45] Although European options have recently begun trading on NYMEX, data was not available on Datastream and is thus not used in the analysis.

[46] The approximation is based on the assumption of an asset with a continuous cost of carry. For futures options, the cost of carry is assumed to be 0.

prices are available. Although the early exercise premium is assumed to be derived from a single factor lognormal model, as it is the implied volatilities rather than the prices that are compared, this approach seems reasonable.

Also, only ATM and OTM options are used in the analysis. There are two reasons for this. The first reason is that it is a convention in the literature as ATM and OTM options tend to be more liquid than ITM options. The other reason is due to Barone-Adesi and Whaley (1987) noting that the American option prices estimated using their formula are likely to have relatively higher errors for ITM options than ATM or OTM options. Additionally, the errors from converting to American option prices are much higher for longer maturity options. Thus only options with up to 8 months to maturity are considered. When filtering the data set taking account of all of the above considerations, the total number of option prices used in the analysis is 15,708.

The European call and put options prices used in the analysis are computed using (4.28) and the characteristic function given by (4.21). For the SV model, we set $\lambda_1 = \lambda_2 = \lambda_3 = 0$ and a numerical solution of (4.25) for computing the characteristic function is used to price options. The numerical solution is obtained in the same manner as for futures prices using a Runge-Kutta method with the *ode45* function in Matlab. For the jump-diffusion and Gaussian models, the solution for the characteristic function of (4.21) is known with $N(\tau) = 0$ and the solution of $M(\tau)$ given by (4.41) in Appendix 4.A. The jump-diffusion model also requires that we set $\rho_{12} = \rho_{13} = \rho_{23} = 0$ and for the Gaussian model, we need to set $\lambda_1 = \lambda_2 = \lambda_3 = 0$ in (4.41). To compute the option prices, we need to evaluate the integrals in (4.28). This is also evaluated numerically using adaptive Gaussian-Kronrod quadrature which is an extension of Gaussian-Legendre quadrature. This is implemented in Matlab using the *quadgk* function which uses adaptive quadrature

based on a Gauss-Kronrod pair ($15^{\text{th}}$ and $7^{\text{th}}$ order formulas). The integrals are also truncated at 1,000 to speed up the calculations[47].

Additionally, the option price equations require the price of zero-coupon bonds (ZCB). To estimate the price of a zero-coupon bond, the yield curve was estimated using a Nelson and Siegel (1987) scheme and applying the procedure to LIBOR and swap data of up to 2 years to maturity. The yields for each option maturity is then interpolated from the yield curves at each time-step to estimate the ZCB prices. Also, it should be noted that the parameters used to price the options are inferred from futures price data only. This means that the estimates not only give an indication of how well each model performs in pricing options, but also whether the information contained in futures prices can be used to accurately estimate options prices.

Of the three specifications considered, only the stochastic volatility model provides additional flexibility when fitting implied volatility surfaces over a range of maturities. Although the JD model should be able to match implied volatility curves for single maturities well, this is generally not the case over a range of maturities and when fitting curves over time. This is because the Poisson processes of the jump component have stationary increments which makes it difficult to match implied volatilities for different sets of maturities. Also, both the jump intensity of the Poisson processes and the volatility of the diffusion component are assumed to be constant which means it is difficult for the model to match implied volatility curves over time. Hence, the SV model should be expected to perform much better across the cross-section of option maturities and strikes for the entire time-series.

[47] Although higher accuracy can be achieved by integrating to a higher value, the differences are negligible when integrating to 10,000 for instance. Hence, the integrals are truncated at 1,000 when calculating option prices.

Table 4.5 reports the root mean-squared errors (RMSE) of the implied volatilities for each model with the columns ordered by strike and rows ordered by option maturity. The results demonstrate that there is little distinguishing the models with no model clearly outperforming any other model over every maturity and moneyness value. If we compute the RMSE for all options together, then it seems that the jump-diffusion model performed the best overall with an RMSE of 3.99% followed by the SV model with 4.20% and 4.71% for the Gaussian model. These results are supported by Table 4.5, where the Gaussian model prices performed the worst for most option categories, although it priced options with moneyness $< 0.9$ the best. Also, although the SV model priced some options well, it performed poorly for a number of other options too. The JD model, however, performed fairly consistently over the entire range although not significantly better than any other models. The errors for all models also tend to be quite high with most option errors being above 3%. It is clear from these results that the implied volatility curves are not fitted well using only futures prices for any of the models including the SV model.

There are a few possible reasons for the poor performance of the models in pricing options when using only futures prices. One possible reason is that option prices generated using futures price data is akin to using historical volatility to price options in equity markets using the Black-Scholes-Merton (BS-M) model. The poor performance when using historical volatility in the BS-M model could possibly be due to the BS-M model being mis-specified and unable to accurately price options. This is a likely possibility in this case, given that none of the models were able to accurately recover option implied volatilities over the entire sample.

Another possibility is that the crude oil futures market exhibits USV. This was mentioned in the previous section where in this case, it means there are risks in the options market not spanned by futures contracts. A weakness of the modelling

**Table 4.5**

RMSE on Option Prices using Futures Data Only

This table reports the option root mean-squared errors (RMSE) where the errors are defined as the difference between the actual implied volatilities and the model implied volatilities reported in percentage points. Options in the table are grouped by moneyness and maturity. The sample period is from 30 August 2006 - 1 August 2008. The option prices for each model were computed assuming parameters estimated from futures data only using (4.25) and (4.28).

| | Moneyness | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Maturity | 0.80 | 0.85 | 0.90 | 0.95 | 1.00 | 1.05 | 1.10 | 1.15 | 1.20 |
| Gaussian | | | | | | | | | |
| 0-1 mth | 9.29 | 4.60 | 4.46 | 6.40 | 6.73 | 6.43 | 5.60 | 4.86 | 6.09 |
| 1-2 mth | 3.88 | 4.02 | 5.06 | 6.25 | 6.27 | 5.94 | 5.38 | 4.90 | 4.36 |
| 2-3 mth | 2.83 | 3.68 | 4.59 | 5.69 | 5.55 | 5.02 | 4.95 | 4.46 | 4.30 |
| 3-4 mth | 2.86 | 3.20 | 4.84 | 4.97 | 5.11 | 4.74 | 4.69 | 4.45 | 4.08 |
| 4-5 mth | 2.74 | 2.96 | 3.88 | 4.78 | 4.92 | 4.61 | 4.48 | 4.46 | 4.04 |
| 5-6 mth | 2.04 | 2.76 | 2.94 | 3.92 | 4.43 | 4.76 | 4.24 | 4.20 | 3.88 |
| 6-7 mth | 1.88 | 2.41 | 3.29 | 4.03 | 4.22 | 3.99 | 4.81 | 4.10 | 3.77 |
| 7-8 mth | 1.98 | 1.56 | 2.47 | 3.03 | 4.00 | 3.78 | 3.49 | 3.51 | 4.32 |
| JD | | | | | | | | | |
| 0-1 mth | 10.45 | 5.63 | 4.20 | 5.55 | 5.59 | 5.33 | 4.95 | 5.96 | 8.73 |
| 1-2 mth | 5.06 | 4.20 | 4.59 | 5.50 | 5.15 | 5.02 | 4.55 | 4.60 | 4.79 |
| 2-3 mth | 3.60 | 3.81 | 4.26 | 5.16 | 4.42 | 3.66 | 3.94 | 3.46 | 4.02 |
| 3-4 mth | 3.37 | 2.80 | 4.68 | 4.52 | 3.97 | 3.04 | 3.24 | 3.16 | 3.40 |
| 4-5 mth | 3.11 | 2.58 | 3.08 | 4.25 | 3.91 | 2.91 | 2.83 | 2.84 | 2.84 |
| 5-6 mth | 2.58 | 2.91 | 2.32 | 2.92 | 3.16 | 2.98 | 2.55 | 2.49 | 2.42 |
| 6-7 mth | 3.01 | 2.92 | 3.41 | 3.64 | 3.09 | 2.46 | 3.15 | 2.57 | 2.21 |
| 7-8 mth | 3.52 | 2.47 | 2.37 | 2.42 | 2.78 | 2.25 | 2.08 | 2.13 | 2.89 |
| SV | | | | | | | | | |
| 0-1 mth | 12.48 | 8.26 | 5.98 | 5.70 | 5.40 | 5.35 | 5.32 | 6.73 | 8.80 |
| 1-2 mth | 6.66 | 5.28 | 4.90 | 5.27 | 4.94 | 4.78 | 4.30 | 4.30 | 4.59 |
| 2-3 mth | 4.00 | 3.88 | 4.18 | 4.94 | 4.16 | 3.47 | 3.71 | 3.35 | 3.69 |
| 3-4 mth | 3.13 | 2.78 | 4.23 | 4.19 | 3.98 | 3.63 | 3.54 | 3.23 | 3.27 |
| 4-5 mth | 2.68 | 2.48 | 3.44 | 4.15 | 4.14 | 3.53 | 3.44 | 3.45 | 3.13 |
| 5-6 mth | 1.97 | 2.62 | 2.42 | 3.56 | 3.96 | 4.43 | 3.49 | 3.47 | 3.20 |
| 6-7 mth | 1.84 | 2.26 | 3.02 | 3.77 | 3.99 | 3.66 | 4.81 | 3.93 | 3.36 |
| 7-8 mth | 1.81 | 1.57 | 2.80 | 2.91 | 4.06 | 3.65 | 3.34 | 3.35 | 4.49 |

approach considered in this study is that it cannot accommodate USV. The model in Hughen (2010) allows for USV where it is shown that if certain parameter restrictions are applied to a suitably specified model, the model can exhibit USV. If the model in this study is reparameterized, then USV could possibly be accommodated with appropriate parameter restrictions. However this would not be possible without substantially altering the model dynamics and changing the interpretation of the state variables. Additionally, a few studies in the term structure literature found that restricting affine term structure models to accommodate USV results in worse option pricing performance than the unrestricted versions[48].

An alternative approach is considered by Trolle and Schwartz (2009) which does not require parameter restrictions to accommodate USV. In their approach, they price commodity derivatives using a HJM framework which models forward curve dynamics rather than spot price dynamics. The approach is appealing as it allows for USV naturally whilst still allowing tractable solutions for derivatives prices. However, the approach does come with some tradeoffs in flexibility compared to the approach in this study and other studies which directly model spot price dynamics. For instance, it is restricted in the choice of volatility structures of the forward cost of carry which allow Markovian dynamics of the spot price. These restrictions mean that the model may not be able to provide realistic dynamics for spot prices. It also makes it difficult to allow for jumps in prices which is easily accommodated under the modelling approach in this study.

Despite the possible presence of USV, this does not imply that option pricing performance of the models is adversely affected. Rather it means that the risk-neutral dynamics of the volatility process cannot be accurately inferred from futures prices

[48]For instance, Bibkov and Chernov (2009) estimated two stochastic volatility models to price eurodollar options, one unrestricted and one restricted to allow for USV. The USV model was found to perform significantly worse than the unrestricted model when pricing the options.

which results in poor option pricing performance. In the next section, the models are calibrated to fit option prices directly which should give a better gauge of the option pricing performance of the SVJ model and its nested specifications.

## 4.4 Estimation using Futures and Options Data

Given the poor options pricing performance in the previous section when using only futures prices, this section looks at how well the models price options using a more direct approach. This is achieved by using both futures and options data to estimate the model parameters. The estimation is conducted using the options data of the previous section from 30 August 2006 - 1 August 2008 as well as futures data spanning the same dates. The futures maturity contracts consist of the same maturities as the previous section with maturities closest to 1, 3, 6, 9, 12, 18 and 24 months. However, only a subset of the options strikes and maturities is used as estimation is quite computationally intensive due to having to calculate option prices using Fourier inversion.

The options are chosen from options on data available on the six closest to maturity futures contracts. For the two closest maturity options, options with moneyness closest to 0.9, 0.95, 1, 1.05 and 1.1 are chosen to ensure that the implied volatility smile is recovered. For longer maturities, only ATM options are chosen to reduce the computational burden and only up to 6 months to maturity. This gives around 14 option prices for each time step to be used for estimation in addition to the 7 futures prices. This gives a maximum total cross-section of around 21 futures and options prices for each daily observation with a total of 482 daily observations.

As options prices are nonlinear functions of the stochastic volatility state variable, $V_t$, the measurement equation of the corresponding state space model will also be nonlinear. This means that applying the standard linear Kalman filter when option prices are present is not appropriate. An alternative is the Unscented Kalman Filter (UKF) introduced by Julier and Uhlmann (1997) for estimating nonlinear state space models. The procedure has been used to estimate models using option prices by Carr and Wu (2007) and Bakshi et al. (2008). It has been found to be superior to the extended Kalman filter (EKF) for estimating nonlinear models whilst having comparable computational efficiency. The UKF also does not require computing Jacobians which is necessary when using the EKF, a generally difficult task for option prices (or adds to computational complexity if numerically obtained). The procedure also has the advantage that as option prices for some maturities were not available for a number of dates, the Kalman filter recursions are still valid and poses no problem in terms of estimation.

Similar to the previous section, it is the implied volatilities of the option prices rather than the actual options prices which are used for estimation. However, as computing options prices is quite computationally expensive[49], computing implied volatilities by inverting option prices is not practical. A much faster alternative considered by Carr and Wu (2007), Bakshi et al. (2008) and Trolle and Schwartz (2009) is to divide the option prices by the Black (1976) vega at the actual option price. As described in Carr and Wu (2007), this essentially converts the options prices to the implied volatility space via a linear approximation. Dividing the true and estimated option price by the vega gives an approximation to the difference in implied volatilities for the two option prices[50]. This quantity is used in the UKF

[49]Evaluating options prices requires computing the integrals for Fourier inversion and the ODE's in (4.25) which must be solved using numerical methods. This requires multiple evaluations to estimate both the solutions to the ODE's and to compute the integrals resulting in a high computational cost.

[50]This is shown by Trolle and Schwartz (2009) by using a first-order Taylor series approximation to the estimated option price. The explanation is as follows. Assume the call option price, $C$,

procedure when computing the model errors.

The state space model for estimating the parameters is similar to the one considered previously in (4.30) - (4.32). However, there are a few differences. The first difference is that the state variables are assumed to be Gaussian even though jumps and stochastic volatility are specified. The main reason for this assumption is that efficient methods for filtering the model such as MCMC or particle filter methods would be very computationally demanding when also computing option prices. This assumption allows application of the standard Kalman filter recursions which are computationally efficient. Also, the main focus in this section is the pricing performance of the models rather than estimating the impact of jumps[51]. Hence, although this assumption means the maximum likelihood estimator is biased, it is still possible to determine the pricing performance of the different model specifications. Given this assumption, the state transition equation is assumed to be

$$X_t = X_{t-1} + (\mu + \mu_j + \kappa X_{t-1})\Delta t + \eta_t, \quad \eta_t \sim N(0, Q_t), \tag{4.34}$$

where $\mu$ and $\kappa$ are the same as for (4.30), $\mu_j = \mathbb{E}(\sum_{i=N_{t-1}}^{N_t} J_i)$ and $Q_t = V_{t-1}\Delta t\Sigma\Sigma' + \text{Var}(\sum_{i=N_{t-1}}^{N_t} J_i)$. Whilst the dependence on the volatility factor is the same as previously, the mean and variance of the jump terms are also included in the equation, yet still assuming that $\eta_t$ is Gaussian.

The other difference in the estimation approach is that it is assumed that the vari-

is a function of its Black (1976) implied volatility, $\sigma$, i.e. $C = f(\sigma)$. Denote the actual option price by $C_0$, its implied volatility by $\sigma_0$ and its vega by $\mathcal{V} = \frac{\partial C}{\partial \sigma}|_{\sigma=\sigma_0}$. Then a first-order Taylor approximation around the estimated option price $C_1$ with corresponding implied volatility, $\sigma_1$ is

$$C_1 \approx C_0 + \mathcal{V}(\sigma_1 - \sigma_0) \Leftrightarrow \sigma_0 - \sigma_1 = (C_0 - C_1)/\mathcal{V}. \tag{4.33}$$

Hence the difference in implied volatilities for the two sets of options prices can be approximated by dividing the difference in the option prices by the Black (1976) vegas.

[51] More efficient estimation of the model and the dynamics of prices under stochastic volatility and jumps is left for future research consideration.

ance of the measurement errors for all futures contracts is the same and the variance of the measurement errors for all options implied volatilities is the same. The main reason for this is again for computational reasons as it means that less parameters need to be estimated. It is also due to options for some strikes and maturities not being available for each observation and it would be impractical to have a different standard deviation for each option as there are generally over 10 different options estimated at any given observation. The measurement equation under this assumption is given by

$$Y_t = h(X_t) + \epsilon_t, N(0, H_t), \tag{4.35}$$

where $h(\cdot)$ maps the state variables to the vector of futures and options prices given by $Y_t \in \mathbb{R}^{(N_{f,t}+N_{o,t})}$, $\epsilon_t \in \mathbb{R}^{(N_{f,t}+N_{o,t})}$ are serially uncorrelated $\mathrm{N}(0, H_t)$ random variables and $N_{f,t}$ and $N_{o,t}$ denote the number of observed futures and options prices at time $t$ respectively. Furthermore, the measurement covariance matrix $H_t$ is assumed to be diagonal with the standard deviation of the measurement errors for the futures prices denoted by $s_{\mathrm{f}}$ and for the options implied volatilities denoted by $s_{\mathrm{o}}$.

**The Unscented Kalman Filter**

Generally, the EKF is applied when the observations are a nonlinear function of the state variables. However, the EKF has been found to perform poorly when the nonlinearities are severe as it relies on a linear approximation based on only a first order Taylor approximation of the nonlinear functions. The UKF was introduced by Julier and Uhlmann (1997) to address some of the issues with the EKF by retaining the nonlinear relationships of the models whilst using a Gaussian approximation to

the distribution of the state variable.

The state distribution is specified by a set of deterministically chosen sample points called sigma points to capture the true mean and covariance of the state variable. When propogated through the nonlinear filter, the posterior mean and covariance are accurately estimated up to the 2nd order for nonlinearity and errors are only incorporated at 3rd or higher orders. The UKF procedure relies on applying an "unscented transformation" to the state variables enabling the calculation of certain statistics on the variables that undergo a nonlinear transformation. The motivation behind the unscented transform is that "it is easier to estimate a probability distribution than an arbitrary nonlinear function or transformation" (Julier and Uhlmann (2004)). The details of the UKF procedure for this application are given in Appendix 4.C.

Similar to the standard Kalman filter, the log-likelihood function for the UKF can be represented in terms of the prediction error, $v_t$, and its covariance matrix, $F_t$. Assuming the errors are normally distributed, the log-likelihood for the state space model of (4.35) and (4.30) given the parameter set, $\Theta$, is

$$L(\Theta) = -\frac{1}{2}\sum_{t=1}^{T}(N_{f,t} + N_{o,t})\log(2\pi) + |F_t| + v_t F_t^{-1} v_t'. \tag{4.36}$$

For both the JD and SV models, the procedure will be a quasi-maximum likelihood procedure as the transition densities in the state equations rely on a Gaussian approximation to the non-Gaussian models. Whilst it is questionable for either model as to whether the approximations are valid, it can be noted that the UKF allows for some flexibility in fitting the first 4 moments of a distribution and should be sufficient if the non-Gaussian features are not too severe. Although there are no studies confirming whether this is the case for these types of models, Carr and Wu

(2007) used the UKF procedure in an application on currency options where the model included jumps and square-root processes. They found little bias in their results, and there is no reason to believe that a large bias exists in this application given some of the similarities in the model dynamics.

### 4.4.1 Estimation Results

Firstly, we briefly comment on the parameter estimates and then compare the model pricing performance. In particular, we will discuss how well the option implied volatilities are recovered from each of the models. Table 4.6 reports the estimated parameters for each model and the mean of the daily RMSE of both futures and option prices for each model. There are a number of features in the table that are worth commenting on. The results are fairly consistent with the results of Table 4.3 in the previous section. For instance, most parameters are significant at the 5% level apart from some of the risk premium parameters and most parameters have the same signs for the corresponding models. Also, the risk-neutral mean-reversion parameters are quite similar for all models. This result is similar to the previous section which indicates that the futures curve dynamics do not differ much between the models. However, the magnitude of most parameters are largely different from the previous section indicating the different price dynamics of the two samples and when accounting for option prices.

If we consider the two models with jumps, JD and SVJ, it should first be noted that as the jumps are not explicitly detected in the estimation process, only their effect on pricing can be commented on. For the JD model, it appears that both of the short-term jump factors are largely insignificant for pricing futures and options and that only the long-term jump factor matters. However, when stochastic volatility

**Table 4.6**
Parameter Estimates

This table reports the parameter estimates for the short/long term Gaussian, jump-diffusion (JD) and stochastic volatility (SV), and stochastic volatility jump models obtained by maximum likelihood using the unscented Kalman filter. The data consists of futures prices with maturities of approximately 1, 3, 6, 9, 12, 18, and 24 months and European futures call and put option prices with maturities of 1 to 6 months. AIC and BIC denotes the Akaike and Bayesian Information Criterion, S.E. denotes the standard errors of the parameter estimates and $\text{RMSE}_f$ and $\text{RMSE}_o$ denote the mean daily RMSE for futures and option prices respectively. The sample period is from 30 August 2006 - 1 August 2008.

| | Gaussian | | JD | | SV | | SVJ | |
|---|---|---|---|---|---|---|---|---|
| Parameter | Estimate | S.E. | Estimate | S.E. | Estimate | S.E. | Estimate | S.E. |
| $\kappa_1^*$ | 4.0634 | (0.0896) | 4.0585 | (0.0897) | 3.9582 | (0.0936) | 3.8573 | (0.0690) |
| $\alpha_1$ | −0.0629 | (0.0794) | −0.0649 | (0.0996) | 0.1838 | (0.0552) | 0.1507 | (0.5111) |
| $\beta_1$ | 0.7275 | (1.4330) | 0.7575 | (2.2517) | 2.6137 | (1.0383) | 3.0029 | (1.2465) |
| $\sigma_1$ | 0.1544 | (0.0081) | 0.1238 | (0.0062) | 0.2051 | (0.0021) | 0.1405 | (0.0039) |
| $\kappa_2^*$ | 1.1752 | (0.0163) | 1.1747 | (0.0163) | 1.1776 | (0.0179) | 1.1684 | (0.0159) |
| $\alpha_2$ | 0.0073 | (0.0762) | 0.0119 | (0.0480) | 0.0165 | (0.0530) | 0.0308 | (0.0569) |
| $\beta_2$ | −2.0721 | (0.7932) | −2.3632 | (1.1848) | −1.8849 | (1.0675) | −1.9604 | (0.2698) |
| $\sigma_2$ | 0.1878 | (0.0067) | 0.1855 | (0.0101) | 0.3213 | (0.0152) | 0.2655 | (0.0076) |
| $\mu_3$ | 0.2632 | (0.1573) | 0.2595 | (0.1723) | 0.1949 | (0.1795) | 0.1781 | (0.1906) |
| $\mu_3^*$ | −0.0369 | (0.0011) | −0.0208 | (0.0037) | −0.0329 | (0.0007) | −0.0258 | (0.0007) |
| $\sigma_3$ | 0.2174 | (0.0045) | 0.1239 | (0.0297) | 0.4200 | (0.0049) | 0.3882 | (0.0088) |
| $\kappa_v^*$ | – | – | – | – | 2.3522 | (0.0131) | 3.4643 | (0.0763) |
| $\theta_v^*$ | – | – | – | – | 0.2512 | (0.0014) | 0.1864 | (0.0088) |
| $\theta_v$ | – | – | – | – | 0.8373 | (0.9788) | 0.7045 | (0.2494) |
| $\beta_v$ | – | – | – | – | 0.6510 | (3.3244) | 0.9578 | (1.9524) |
| $\sigma_v$ | – | – | – | – | 1.0849 | (0.0039) | 0.9292 | (0.0255) |
| $\rho_{12}$ | −0.2110 | (0.0512) | −0.2122 | (0.0543) | −0.3785 | (0.0442) | −0.9641 | (0.0072) |
| $\rho_{13}$ | 0.0887 | (0.0516) | 0.1729 | (0.1056) | −0.1891 | (0.0292) | −0.6796 | (0.0327) |
| $\rho_{1v}$ | – | – | – | – | 0.2588 | (0.0017) | 0.4235 | (0.0135) |
| $\rho_{23}$ | 0.2664 | (0.0408) | 0.7387 | (0.3599) | 0.3131 | (0.0214) | 0.4624 | (0.0158) |
| $\rho_{2v}$ | – | – | – | – | −0.0069 | (0.0006) | −0.0470 | (0.0013) |
| $\rho_{3v}$ | – | – | – | – | −0.3469 | (0.0022) | −0.2535 | (0.0001) |
| $\lambda_1$ | – | – | 0.2305 | (4.3433) | – | – | 0.0190 | (0.0040) |
| $\omega_u$ | – | – | 80.0254 | (26.7185) | – | – | 1.2347 | (0.0969) |
| $\lambda_2$ | – | – | 0.3344 | (0.4071) | – | – | 0.7589 | (0.2257) |
| $\omega_d$ | – | – | 12.0349 | (6.3064) | – | – | 8.7235 | (1.3755) |
| $\lambda_3$ | – | – | 3.0642 | (1.4594) | – | – | 0.0872 | (3.4301) |
| $\sigma_j$ | – | – | 0.0821 | (0.0151) | – | – | 0.0093 | (0.3355) |
| $s_\text{f}$ | 0.0016 | (0.0000) | 0.0016 | (0.0000) | 0.0017 | (0.0000) | 0.0017 | (0.0000) |
| $s_\text{o}$ | 0.0494 | (0.0004) | 0.0493 | (0.0004) | 0.0145 | (0.0001) | 0.0142 | (0.0001) |
| Loglik | 24196.81 | | 24213.87 | | 31321.38 | | 31544.18 | |
| AIC | −48361.62 | | −48383.75 | | −62594.76 | | −63028.35 | |
| BIC | −48294.77 | | −48291.83 | | −62494.49 | | −62903.02 | |
| $\text{RMSE}_\text{f}$ | 0.12 | | 0.12 | | 0.12 | | 0.12 | |
| $\text{RMSE}_\text{o}$ | 4.36 | | 4.32 | | 1.28 | | 1.22 | |

and jumps are introduced, the reverse seems to hold true as only the jumps in the two short-term factors matter and long-term jumps are insignificant. To explain this, it should be noted that from the previous section it was found that jumps have little effect on futures pricing performance so that it is option prices that are mainly affected by introducing jumps. For the JD model, implied volatility for a particular strike and moneyness is constant over time given the model parameters. This could mean that the long-term jump factor is able to better match the average implied volatility smile than the two short-term jump factors. However, for the SVJ model, the implied volatility smiles are time-varying due to the volatility factor and the average smile is accounted for by the volatility factor. In this case, the normally distributed long-term jumps are mainly useful for matching the steep volatility smiles observed in shorter maturity options. However, on inspection of the volatility smiles, it appears that the volatility smiles are skewed. The jumps in the two short-term factors are able to introduce some skew in the implied volatility smile if either the jump intensity or jump sizes differ. Given that the jump parameters are substantially different and significant for the two short-term factors, it seems that they are able to capture the skew as well as match implied volatility smiles at shorter maturities. This renders jumps in the long-term factor obsolete in option pricing when stochastic volatility is also present.

For the stochastic volatility factor, the risk-neutral parameters are highly significant for both the SV and SVJ models. Comparing the parameters between the two models, the mean-reversion of the volatility factor in the SVJ model is higher, whereas the long-term means and volatility-of-volatility parameters are lower. This is to be expected due to the SVJ model accounting for jumps. Hence, as the SV model does not take account of jumps in prices, the persistence in volatility, long-term means and volatility-of-volatility should be inflated. Also, commenting on the risk premium parameters for the volatiltiy factor, the parameters are mostly

not significant. For instance, $\beta_v$ is not significant for either model and $\theta_v$ is only significant for the SVJ model. However, the values are consistent with the futures only estimation results except this suggest that negative volatility risk premiums are possibly priced in options as well as futures. This is not conclusive though, since the parameters are not significant.

In terms of pricing, there is little difference in the futures pricing performance for all of the models with an error of around 0.12%. This shows that even when pricing options in addition to futures, the futures pricing performance of the models does not differ much. This is consistent with our previous observation that the risk-neutral mean-reversion parameters are similar. However, there is clearly a difference in the option pricing performance as both the Gaussian and JD models have a RMSE of over 4% whereas the SV and SVJ models have a RMSE of 1.22% and 1.28% respectively. These results are more in line with the expected performance of the models. This shows that introducing jumps only marginally improves option pricing fit over the time-series. This is because when the jump intensities are assumed constant, the model does not have the flexibility to account for time-varying implied volatilities. However, incorporating stochastic volatility results in a large improvement in option price fits as it allows for more flexibility when fitting the implied volatility surface across both the maturity and moneyness space.

As these errors were only for a subset of the total options data, an inspection of the option pricing performance for the full data set is also conducted. Figure 4.2 shows the mean errors vs. moneyness of each of the models for 1,3 and 6 month options. The plot gives a visual indication of whether any bias can be found in the implied volatility fits. The plot shows that 0-1 month options exhibit the most bias for each model, although the SVJ has the least amount of bias with mean errors closest to zero. However, all of the other model specifications underestimate implied volatility

for 0-1 month options with mean errors substantially above zero and a marked 'smile' shape can be observed indicating a bias for deep OTM options. This clearly demonstrates the inadequacy of the Gaussian and JD models to match implied volatility smiles for short-term options with mean errors above 5% for options with moneyness of 0.8 or 1.2. The bias decreases with increasing maturity as the mean errors are mostly within 2% for all of the models. The bias also appears to be higher for deep OTM put options. This can be seen by the higher mean errors for put options with a moneyness of 0.8 compared to call options with a moneyness of 1.2. Interestingly, the SVJ model appears to price 6 month options worse than the SV model. This suggests that the SVJ model sacrifices some of its overall fit to better capture the implied volatility smile at shorter maturities.

We now consider the RMSE of implied volatilities which are reported in a similar manner as the previous section. The RMSE for each model is reported in Table 4.7 with the errors grouped by moneyness and maturity. The option prices used to compute the errors for each model are based on the parameter estimates in Table 4.6. The pricing errors are reported on the full-sample used in Table 4.5 and includes the SVJ model in the analysis. Of the three models analysed in Table 4.5, only the SV model demonstrates a substantial decrease in the option pricing errors. The errors in the implied volatility estimates for the SV model are less than half the size of the errors when estimating the model using futures data only. Although the Gaussian and jump diffusion models demonstrate some improvement in option pricing performance, as described previously, they are not able to replicate the implied volatility curves across both the moneyness and maturity dimensions. This is especially true for the Gaussian model where the implied volatility curves are flat for each maturity resulting in the worst performance for deep OTM options.

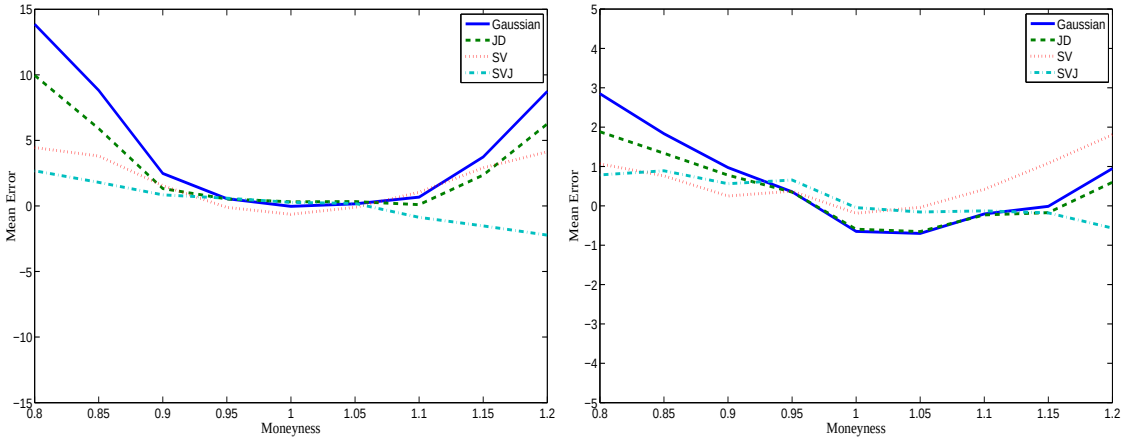
If we consider the overall performance, it is clear that the SVJ model is the best

**Table 4.7**

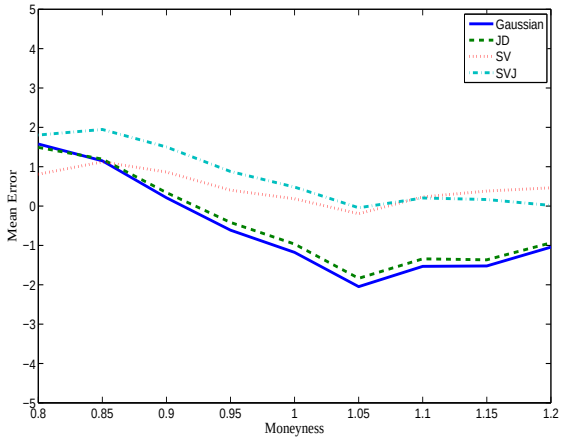
RMSE on Option Prices using Futures and Options Data

This table reports the option root mean-squared errors (RMSE) where the errors are defined as the difference between the actual implied volatilities and the model implied volatilities reported in percentage points. Options in the table are grouped by moneyness and maturity. The sample period is from 30 August 2006 - 1 August 2008. The option prices for each model were computed assuming parameters estimated from futures and options data using (4.25) and (4.28).

| | Moneyness | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Maturity | 0.80 | 0.85 | 0.90 | 0.95 | 1.00 | 1.05 | 1.10 | 1.15 | 1.20 |
| Gaussian | | | | | | | | | |
| 0-1 mth | 7.49 | 6.69 | 4.72 | 5.55 | 5.55 | 5.32 | 5.04 | 6.27 | 9.17 |
| 1-2 mth | 5.44 | 4.63 | 4.67 | 5.51 | 5.11 | 5.02 | 4.57 | 4.66 | 4.87 |
| 2-3 mth | 3.98 | 4.00 | 4.32 | 5.17 | 4.40 | 3.65 | 3.94 | 3.47 | 4.01 |
| 3-4 mth | 3.52 | 2.84 | 4.70 | 4.53 | 3.99 | 3.08 | 3.29 | 3.20 | 3.39 |
| 4-5 mth | 3.14 | 2.58 | 3.08 | 4.25 | 3.95 | 3.01 | 2.93 | 2.94 | 2.88 |
| 5-6 mth | 2.53 | 2.85 | 2.30 | 2.97 | 3.24 | 3.15 | 2.70 | 2.65 | 2.54 |
| 6-7 mth | 2.87 | 2.80 | 3.32 | 3.62 | 3.18 | 2.62 | 3.36 | 2.75 | 2.39 |
| 7-8 mth | 3.29 | 2.25 | 2.29 | 2.40 | 2.92 | 2.46 | 2.25 | 2.30 | 3.10 |
| JD | | | | | | | | | |
| 0-1 mth | 5.46 | 4.66 | 4.14 | 5.55 | 5.55 | 5.32 | 4.97 | 5.29 | 6.77 |
| 1-2 mth | 4.04 | 4.10 | 4.59 | 5.51 | 5.10 | 5.03 | 4.56 | 4.53 | 4.43 |
| 2-3 mth | 3.36 | 3.80 | 4.28 | 5.17 | 4.39 | 3.64 | 3.94 | 3.46 | 3.93 |
| 3-4 mth | 3.24 | 2.78 | 4.68 | 4.53 | 3.97 | 3.05 | 3.27 | 3.20 | 3.38 |
| 4-5 mth | 3.01 | 2.56 | 3.08 | 4.25 | 3.91 | 2.94 | 2.87 | 2.91 | 2.88 |
| 5-6 mth | 2.48 | 2.87 | 2.31 | 2.92 | 3.17 | 3.02 | 2.60 | 2.57 | 2.50 |
| 6-7 mth | 2.91 | 2.88 | 3.40 | 3.63 | 3.10 | 2.48 | 3.20 | 2.63 | 2.29 |
| 7-8 mth | 3.44 | 2.44 | 2.37 | 2.42 | 2.78 | 2.27 | 2.10 | 2.16 | 2.95 |
| SV | | | | | | | | | |
| 0-1 mth | 2.43 | 3.18 | 2.00 | 1.16 | 1.37 | 1.25 | 2.16 | 3.95 | 5.93 |
| 1-2 mth | 2.53 | 1.80 | 1.02 | 1.71 | 1.04 | 1.05 | 1.43 | 2.20 | 3.13 |
| 2-3 mth | 1.57 | 1.22 | 0.77 | 2.82 | 0.91 | 0.82 | 0.95 | 1.53 | 2.22 |
| 3-4 mth | 1.31 | 1.04 | 1.09 | 2.01 | 1.08 | 1.06 | 0.95 | 1.22 | 1.74 |
| 4-5 mth | 1.16 | 1.05 | 2.34 | 1.78 | 1.32 | 1.18 | 0.94 | 0.98 | 1.21 |
| 5-6 mth | 1.13 | 2.45 | 1.79 | 1.61 | 1.29 | 1.44 | 1.03 | 0.94 | 1.09 |
| 6-7 mth | 1.65 | 1.96 | 2.04 | 2.05 | 1.54 | 1.19 | 1.33 | 1.08 | 0.86 |
| 7-8 mth | 1.85 | 1.87 | 1.66 | 1.62 | 1.52 | 1.13 | 0.98 | 0.89 | 1.24 |
| SVJ | | | | | | | | | |
| 0-1 mth | 1.44 | 1.59 | 1.60 | 1.38 | 1.27 | 1.29 | 1.97 | 2.66 | 3.38 |
| 1-2 mth | 1.46 | 1.42 | 1.11 | 1.77 | 0.97 | 1.06 | 1.27 | 1.77 | 2.87 |
| 2-3 mth | 1.48 | 1.33 | 0.91 | 2.84 | 0.88 | 0.84 | 0.85 | 1.10 | 1.49 |
| 3-4 mth | 1.50 | 1.28 | 1.28 | 2.04 | 1.07 | 1.07 | 0.89 | 0.95 | 1.38 |
| 4-5 mth | 1.62 | 1.39 | 2.46 | 1.84 | 1.30 | 1.15 | 0.91 | 0.88 | 1.01 |
| 5-6 mth | 1.96 | 2.89 | 2.16 | 1.72 | 1.31 | 1.40 | 1.02 | 0.87 | 0.99 |
| 6-7 mth | 2.60 | 2.67 | 2.53 | 2.31 | 1.65 | 1.25 | 1.29 | 1.06 | 0.82 |
| 7-8 mth | 3.05 | 2.75 | 2.32 | 2.06 | 1.69 | 1.31 | 1.18 | 1.00 | 1.25 |




**(a)** 1 Month **(b)** 3 Month




**(c)** 6 Month

**Figure 4.2**
Mean option errors for 1 month, 3 month and 6 month options. Each plot shows errors for the Gaussian (solid line), JD (dashed line), SV (dotted line) and SVJ (dash-dot line) models.

performing model. It demonstrates the best overall fit of option implied volatilities across maturities and moneyness and substantially outperforms the other models for near maturity options. The worst RMSE of the SVJ model is 3.38% for OTM 1-month call options with a moneyness of 1.2. This is much better than the SV model with an RMSE of 5.93% for the same group of options and is better than the RMSE of most option categories for the Gaussian and JD models. However, although the overall performance of the SV was worse than the SVJ model, it does fit some option categories better than the SVJ model which suggests its performance is also quite good.

We also plot the volatility estimated under the SV amd SVJ models and compare it to the nearby futures price volatility in Figure 4.3. The approach is similar to the previous section where the nearby futures price volatility is estimated using the discrete-time SV model which is labelled as historical volatility in the plot. Interestingly, the volatility plots of both the SV and SVJ models track the nearby futures volatility quite closely which indicates that at least for this sample, volatility is largely spanned by futures prices. However, there is some divergence of the two models from the historical volatility line from around March 2008 onwards. This was during a time of high uncertainty in financial markets leading up to the peak of the GFC in September 2009. As mentioned previously, a possible reason for the divergence is that the crude oil options market exhibits USV and that the options market accounts for risk during this period not accounted for by futures contracts.

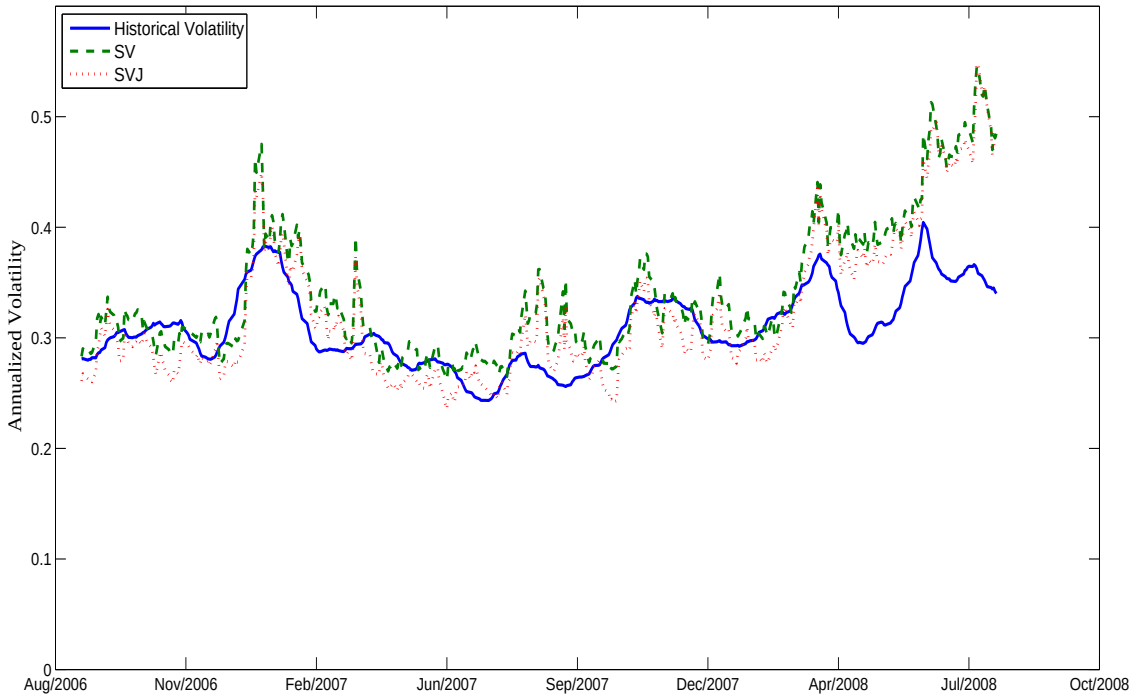



**Figure 4.3**
Plot of historical volatility and estimated volatility for the SV and SVJ models. Volatility is measured on the nearby futures contract.

This is a possible reason for the poor option pricing performance of the SV model when using only futures prices to estimate the model parameters compared to using

both futures and option prices. Although this does suggest that USV may be present in the crude oil futures and options markets, this does not necessarily imply that futures and option pricing performance will be improved. Trolle and Schwartz (2009) found evidence for USV in the crude oil options market, although their study only considered a stochastic volatility model without accounting for jumps. Although they did not deem jumps as important for pricing options, these results suggest that jumps are necessary for matching the implied volatility smiles of short-term options. It is also a possibility that a model with jumps can explain the behaviour of spot prices which an SV model cannot, such as infrequent but large price moves. Whilst USV could possibly improve both futures and option pricing performance, as the main focus in this study is on the impact of jumps and stochastic volatility on pricing dynamics, the issue is left for future research consideration.

The results of this section show that although the Gaussian and JD models are sufficient when pricing futures prices, this is not the case for options. They clearly do not have the flexibility to fit the option implied volatility surface when considering both the cross-sectional and time-series fit. This suggests that stochastic volatility is necessary when pricing options, whilst jumps are also important but only in conjunction with stochastic volatility. This is especially the case when attempting to fit volatility smiles for options close to maturity. These results suggest that crude oil option prices exhibit similar features to other financial options markets. Hence, some of the modelling approaches introduced in those markets have potential to improve modelling of commodities.

## 4.5 Conclusion

This chapter considered an extension to the three-factor short-long diffusion model with jumps and stochastic volatility. The purpose of the extensions was to capture some of the empirical characteristics evident in commodity markets not explained by a Gaussian model and to fill a gap in the literature. This includes infrequent but large movements in spot prices, time-varying volatility and the volatility clustering effect. The model in this study introduced jumps which allow for skewness in returns, jumps which affect short-term futures prices more than long term prices, and stochastic volatility. We also derived semi-closed form solutions for futures and European option prices for the model.

Using futures price data on WTI crude oil traded on NYMEX, this study conducted an empirical analysis of the model and its nested specifications. The results showed that incorporating either jumps or stochastic volatility was able to explain the dynamics of the crude oil price process better than a Gaussian model. However when fitting the futures term structure, there is little improvement in performance as futures prices rely only on the first moment of the stock price distribution of which jumps and stochastic volatility add little.

In order to gauge the effects of jumps or stochastic volatility on higher moment effects, a study of the model's option pricing performance is conducted. Incorporating either jumps or stochastic volatility was shown to improve on the option pricing fits over a Gaussian model. However, it is clear that using futures prices alone to estimate the models is insufficient when it comes to pricing options. This suggests that options possibly contain information not accounted for by futures prices.

When estimating the model using both futures and options prices, the findings suggest that there are clear improvements in option pricing performance when stochastic volatility and jumps are added to the Gaussian model. Although stochastic volatility is required to match option prices over the moneyness and maturity space as well as over time, jumps substantially improve the fit of short-term options. These results are similar to empirical studies of financial options and demonstrates that the same holds in the crude oil options market.

# Appendices

## Appendix 4.A Futures and Option Prices of the Jump-Diffusion Model

For the jump-diffusion model with no stochastic volatility, it is assumed that $V_t$ is constant and set to 1 in (4.1) and (4.6). Under the nested model, the risk-neutral dynamics are

$$\begin{aligned} dX_{1,t} &= (-\alpha_1 - \kappa_1^* X_{1,t})dt + \sigma_1 dZ_{1,t}^* + J_{1,t}dN_{1,t} - \lambda_1(\nu_1(1)-1)dt, \\ dX_{2,t} &= (-\alpha_2 - \kappa_2^* X_{2,t})dt + \sigma_2 dZ_{2,t}^* + J_{2,t}dN_{2,t} - \lambda_2(\nu_2(1)-1)dt, \\ dX_{3,t} &= (\mu_3 - \alpha_3)dt + \sigma_3 dZ_{3,t}^* + J_{3,t}dN_{3,t} - \lambda_3(\nu_3(1)-1)dt. \end{aligned} \tag{4.37}$$

For the futures price, given the PDE for the futures price in (4.13), if $V_t$ is assumed constant and set to 1 (or equivalently setting $V_0 = 1$, $\theta_v = 1$ and $\sigma_v = 0$), then the futures price can be seen as the solution to the following system of ODE's,

$$B_1'(\tau) = -\kappa_1 B_1(\tau), \tag{4.38a}$$

$$B_2'(\tau) = -\kappa_2 B_2(\tau), \tag{4.38b}$$

$$B_3'(\tau) = 0, \tag{4.38c}$$

$$\begin{aligned} A'(\tau) = &-\alpha_1 B_1(\tau) - \alpha_2 B_2(\tau) + (\mu_3 - \alpha_3)B_3(\tau) + \frac{1}{2}\Big[B_1(\tau)^2\sigma_1^2 + B_2(\tau)^2\sigma_2^2 \\ &+ B_3(\tau)^2\sigma_3^2 + 2B_1(\tau)B_2(\tau)\sigma_1\sigma_2\rho_{12} + 2B_1(\tau)B_3(\tau)\sigma_1\sigma_3\rho_{13} \\ &+ 2B_2(\tau)B_3(\tau)\sigma_2\sigma_3\rho_{23}\Big] + \lambda_1\big[\nu_1(B_1(\tau)) - 1 - B_1(\tau)(\nu_1(1)-1)\big] \\ &+ \lambda_2\big[\nu_2(B_2(\tau)) - 1 - B_2(\tau)(\nu_2(1)-1)\big] \\ &+ \lambda_3\big[\nu_3(B_3(\tau)) - 1 - B_3(\tau)(\nu_3(1)-1)\big]. \end{aligned} \tag{4.38d}$$

Solving the system of ODE's leads to the following expression for the futures price

$$F(t,T) = \exp\big\{A(\tau) + B(\tau)\cdot X_t\big\}, \tag{4.39}$$

where

$$\begin{aligned}
&B_1(\tau) = e^{-\kappa_1^*\tau},\ B_2(\tau) = e^{-\kappa_2^*\tau},\ B_3(\tau) = 1,\\
&A(\tau) = -\frac{\alpha_1}{\kappa_1^*}\big(1-e^{-\kappa_1^*\tau}\big) - \frac{\alpha_2}{\kappa_2^*}\big(1-e^{-\kappa_2^*\tau}\big) + (\mu_3-\alpha_3)\tau + \frac{1}{2}\Bigg[(1-e^{-2\kappa_1^*\tau})\frac{\sigma_1^2}{2\kappa_1^*}\\
&\quad + \big(1-e^{-2\kappa_1^*\tau}\big)\frac{\sigma_1^2}{2\kappa_1^*} + \sigma_3^2\tau + \big(1-e^{-(\kappa_1^*+\kappa_2^*)\tau}\big)\frac{2\rho_{12}\sigma_1\sigma_2}{\kappa_1^*+\kappa_2^*}\\
&\quad + \big(1-e^{-\kappa_1^*\tau}\big)\frac{2\sigma_1\sigma_3\rho_{13}}{\kappa_1^*} + \big(1-e^{-\kappa_2^*\tau}\big)\frac{2\sigma_2\sigma_3\rho_{23}}{\kappa_2^*}\Bigg]\\
&\quad + \frac{\lambda_1}{\kappa_1^*}\Bigg[\log\Big(\frac{\omega_u - e^{-\kappa_1^*\tau}}{\omega_u-1}\Big) - \frac{1-e^{-\kappa_1^*\tau}}{\omega_u-1}\Bigg] + \frac{\lambda_2}{\kappa_2^*}\Bigg[\log\Big(\frac{\omega_d + e^{-\kappa_2^*\tau}}{\omega_d+1}\Big) + \frac{1-e^{-\kappa_2^*\tau}}{\omega_d+1}\Bigg].
\end{aligned}$$

Similarly for the characteristic function given by (4.21) and (4.22) with $V_t = 1$ and $N(\tau) = 0$, $M(\tau)$ is a solution to the following ODE,

$$\begin{aligned}
M'(\tau) = \frac{1}{2}iu(iu-1)\Big(&e^{-2\kappa_1^*\tau_F}\sigma_1^2 + e^{-2\kappa_2^*\tau_F}\sigma_2^2 + \sigma_3^2 + 2e^{-(\kappa_1^*+\kappa_2^*)\tau_F}\sigma_1\sigma_2\rho_{12}\\
&+ 2e^{-\kappa_1^*\tau_F}\sigma_1\sigma_3\rho_{13} + 2e^{-\kappa_2^*\tau_F}\sigma_2\sigma_3\rho_{23}\Big) + \lambda_1\Big(\frac{\omega_u}{\omega_u - iue^{-\kappa_1^*\tau_F}} - 1\Big)\\
&- iu\lambda_1\Big(\frac{\omega_u}{\omega_u - e^{-\kappa_1^*\tau_F}} - 1\Big) + \lambda_2\Big(\frac{\omega_d}{\omega_d + iue^{-\kappa_2^*\tau_F}} - 1\Big)\\
&- iu\lambda_2\Big(\frac{\omega_d}{\omega_d + e^{-\kappa_2^*\tau_F}} - 1\Big) + \lambda_3\Big(e^{-u^2\sigma_j^2/2} - 1\Big) - iu\lambda_3\Big(e^{\sigma_j^2/2} - 1\Big).
\end{aligned} \tag{4.40}$$

This leads to the following solution for $M(\tau)$,

$$\begin{aligned} M(\tau) = \frac{1}{2}iu(iu-1)\Bigg[ & \frac{\big(1-e^{-2\kappa_1^*\tau}\big)e^{-2\kappa_1^*\bar{\tau}}\sigma_1^2}{2\kappa_1^*} + \frac{\big(1-e^{-2\kappa_2^*\tau}\big)e^{-2\kappa_2^*\bar{\tau}}\sigma_2^2}{2\kappa_2^*} + \sigma_3^2\tau \\ & + \frac{2\big(1-e^{-(\kappa_1^*+\kappa_2^*)\tau}\big)e^{-(\kappa_1^*+\kappa_2^*)\bar{\tau}}\rho_{12}\sigma_1\sigma_2}{\kappa_1^*+\kappa_2^*} + \frac{2\big(1-e^{-\kappa_1^*\tau}\big)e^{-\kappa_1^*\bar{\tau}}\sigma_1\sigma_3\rho_{13}}{\kappa_1^*} \\ & + \frac{2\big(1-e^{-\kappa_2^*\tau}\big)e^{-\kappa_2^*\bar{\tau}}\sigma_2\sigma_3\rho_{23}}{\kappa_2^*}\Bigg] + \frac{\lambda_1}{\kappa_1^*}\Bigg[\log\Big(\frac{\omega_u - iue^{-\kappa_1^*(\bar{\tau}+\tau)}}{\omega_u - iue^{-\kappa_1^*\bar{\tau}}}\Big) \\ & - iu\log\Big(\frac{\omega_u - e^{-\kappa_1^*(\bar{\tau}+\tau)}}{\omega_u - e^{-\kappa_1^*\bar{\tau}}}\Big)\Bigg] + \frac{\lambda_2}{\kappa_2^*}\Bigg[\log\Big(\frac{\omega_d + iue^{-\kappa_2^*(\bar{\tau}+\tau)}}{\omega_d + iue^{-\kappa_2^*\bar{\tau}}}\Big) \\ & - iu\log\Big(\frac{\omega_d + e^{-\kappa_2^*(\bar{\tau}+\tau)}}{\omega_d + e^{-\kappa_2^*\bar{\tau}}}\Big)\Bigg] + \lambda_3\Big(e^{-u^2\sigma_j^2/2} - iue^{\sigma_j^2/2} + iu - 1\Big)\tau. \end{aligned} \tag{4.41}$$

European call option prices can then be computed using (4.28).

# Appendix 4.B MCL Estimation of the Jump-Diffusion Model

This appendix describes estimation of the jump-diffusion model with no stochastic volatility. In this section, we assume the state space model given by (4.30) and (4.35) but without volatility factor and without the assumption of normality in the state equation. Hence the state vector is defined by $X_t = (X_{1,t}, X_{2,t}, X_{3,t})'$. The system is linear although incorporating jumps means the state space model is non-Gaussian. If we set $\eta_t = [\Delta Z_t + J_t \cdot q_t]'$, then the state space model can be described alternatively as,

$$\begin{aligned} y_t &= Z_t x_t + d_t + \epsilon_t, \quad \epsilon_t \sim N(0, H_t), \\ x_{t+1} &= T_t x_t + c_t + \eta_t, \quad \eta_t \sim p(\eta_t), \end{aligned} \tag{4.42}$$

for $t = 1, \ldots, T$ where $y_t$ is the vector of observed log-futures prices at time $t$ and $x_t = X_t$. As described in Chapter 2, the Monte Carlo likelihood (MCL) estimation procedure involves using importance sampling techniques to estimate nonlinear, non-Gaussian state space models. The approach here follows Dempster et al. (2008) who applied the MCL procedure to estimate a three-factor jump-diffusion model. Assuming the state space model given by (4.42), Durbin and Koopman (1997) show that the likelihood function can be represented as

$$L(y|\Theta) = L_G(y|\Theta)\mathbb{E}_G\left(\frac{p(x,y)}{q(x,y)}\right), \tag{4.43}$$

where $L(\cdot|\Theta)$ is the true likelihood, $L_G(\cdot|\Theta)$ is the likelihood of the approximating linear Gaussian model, $\Theta$ denotes the parameter set and $\mathbb{E}_G(\cdot)$ is expectation under the approximating model's density. Hence the aim is to estimate this likelihood function. Although the procedure is similar to the application in Chapter 2, there are some differences when adapting the procedure to this model. Firstly, the linear approximating Gaussian model is assumed to have the following state space form,

$$\begin{aligned} y_t &= Z_t x_t + d_t + \epsilon_t, \quad \epsilon_t \sim N(0, H_t), \\ x_{t+1} &= T_t x_t + c_t + \eta_t, \quad \eta_t \sim N(0, Q_t), \end{aligned} \tag{4.44}$$

for $t = 1, \ldots, T$ which is similar to (4.42) except that $\eta_t$ is Gaussian in this equation. Denote the conditional and joint densities of the general state space model by $p(x|y)$ and $p(x, y)$ and the linear Gaussian approximating densities by $q(x|y)$ and $q(x, y)$. Durbin and Koopman (1997, 2001) show that the mode $\hat{x}$ of the above model is the solution to $\partial \log q(x|y)/\partial x = 0$. However, since $q(x|y) = q(x, y) + q(y)$, $\hat{x}$ is also the solution to $\partial q(x, y)/\partial x$ and as $q(x, y)$ generally has a more tractable form than $q(x|y)$, this is the quantity we work with. As further specified in Durbin and Koopman (1997, 2001) or derived from the model above, the linear Gaussian state

space model has the following log joint density

$$\log q(x,y) = \text{constant} - \frac{1}{2}\sum_{t=1}^{T}\left[(x_t - T_t x_{t-1} - c_t)' Q_t^{-1}(x_t - T_t x_{t-1}) + (y_t - Z_t x_t - d_t)' H_t^{-1}(y_t - Z_t x_t - d_t)\right], \tag{4.45}$$

and differentiating with respect to $x_t$, $t = 1, \ldots, T$ and equating to zero gives

$$-Q_t^{-1}(x_t - T_t x_{t-1} - c_t) + i_t T_{t+1}' Q_{t+1}^{-1}(x_{t+1} - T_{t+1} x_t - c_{t+1}) + Z_t' H_t^{-1}(y_t - Z_t x_t) = 0, \tag{4.46}$$

where $i_t = 1$ for $t < T$ and $i_t = 0$ for $t = T$. The solutions to these sets of equations is the conditional mode $\hat{x}$, but as the model is a linear Gaussian state space model, the mode coincides with the mean. Hence the conditional mode can be obtained through an application of the Kalman filter and smoother recursions.

For the non-Gaussian state space model described by (4.42), the log joint density is,

$$\log p(x,y) = \text{constant} - \sum_{t=1}^{T}\left[\log p(\eta_t) + \frac{1}{2}(y_t - Z_t x_t - d_t)' H_t^{-1}(y_t - Z_t x_t - d_t)\right]. \tag{4.47}$$

As with the linear Gaussian model, the mode of $p(x|y)$ is a solution to $\partial p(x,y)/\partial x = 0$. Hence, by taking the partial derivative with respect to $x_t$ and equating to 0, the mode $\hat{x}$ is a solution to

$$\frac{\partial \log p(\eta_t)}{\partial \eta_t} + i_t T_{t+1}' \frac{\partial \log p(\eta_{t+1})}{\partial \eta_{t+1}} + Z_t' H_t^{-1}(y_t - Z_t x_t) = 0, \tag{4.48}$$

where $\eta_t = x_t - T_t x_{t-1} - c_t$, $t = 1, \ldots, T$, $i_t = 1$ for $t < T$ and $i_t = 0$ for $t = T$. As can be seen from the form of both (4.46) and (4.48), the two models will have the

same mode if the following set of equalities hold,

$$-Q_t^{-1}\eta_t = \frac{\partial \log p(\eta_t)}{\partial \eta_t} \quad \text{or} \quad Q_t^{-1} = -\frac{\partial \log p(\eta_t)}{\partial \eta_t}/\eta_t. \tag{4.49}$$

For the above equation to have a unique solution, the state errors must be assumed to be independent. If correlation between each of the state variables is introduced, then the inequalities above will result in a system of equations with more variables to solve than equations. To ensure that there is a unique solution to the system of equations, the correlation between each state variable is set to 0. This means that $\rho_{ij} = 0$ for $i \neq j$ in (4.37) so that $Q_t$ is diagonal.

Given the above, we thus need to compute both the densities and the partial derivatives of the densities for each of the factors. For the exponentially distributed jumps in the two short-term factors, we use the densities derived in Ramezani and Zeng (2007) who use the fact that the sum of independent exponentially distributed random variables are Gamma distributed. This means that if $x_i \sim \exp(\omega_u)$ for $i = 1, \ldots, n$ and $X = \sum_{i=1}^{n} x_i$, then $X \sim \Gamma(n, \omega_u)$ where

$$p(X|n) = \frac{\omega_u^n}{(n-1)!} X^{n-1} e^{-\omega_u X}. \tag{4.50}$$

Similarly, for the negatively distributed jumps, if $-y_i \sim \exp(\omega_d)$ for $i = 1, \ldots, n$ and $Y = \sum_{i=1}^{n} x_i$, then $-z \sim \Gamma(n, \omega_d)$ where

$$p(Y|n) = \frac{\omega_d^n}{(n-1)!} (-Y)^{n-1} e^{\omega_d Y}. \tag{4.51}$$

For each factor, the number of jumps that occur between each observation is denoted by $\Delta N_{i,t} = N_{i,t} - N_{i,t-1}$. For $\eta_{1,t}$, only positive exponentially distributed jumps occur

so an approximation to the density is given by[52]

$$\begin{aligned} p(\eta_{1,t}) &\approx \sum_{i=0}^{\infty} \Pr(N_{1,t} = i) p_1(\eta_{1,t} | \Delta N_{1,t} = i) \\ &= \sum_{i=1}^{\infty} \Pr(\Delta N_{1,t} = i) \frac{\omega_u^i}{(i-1)!} \int_0^{\infty} x^{i-1} e^{-\omega_u x} \phi(\eta_{1,t}; x, \sigma_1^2 \Delta t) dx, \end{aligned} \tag{4.52}$$

where $\phi(x; \mu, \sigma^2)$ denotes the normal pdf with mean $\mu$ and variance $\sigma^2$. Similarly, for $\eta_{2,t}$, the jumps are negative exponentially distributed so it has a pdf given by,

$$\begin{aligned} p(\eta_{2,t}) &\approx \sum_{i=0}^{\infty} \Pr(\Delta N_{2,t} = i) p_2(\eta_{2,t} | N_{2,t} = i) \\ &= \sum_{i=1}^{\infty} \Pr(\Delta N_{2,t} = i) \frac{\omega_d^i}{(i-1)!} \int_0^{\infty} (-x)^{i-1} e^{\omega_d x} \phi(\eta_{2,t}; x, \sigma_2^2 \Delta t) dx. \end{aligned} \tag{4.53}$$

For the long-term factor, the jumps are normally distributed, so the density of $\eta_{3,t}$ is

$$\begin{aligned} p(\eta_{3,t}) &= \sum_{i=0}^{\infty} \Pr(\Delta N_{3,t} = i) p_3(\eta_{3,t} | \Delta N_{3,t} = i) \\ &= \sum_{i=0}^{\infty} \Pr(\Delta N_{3,t} = i) \phi(\eta_{3,t}; i\mu_j, \sigma_3^2 \Delta t + i\sigma_j^2), \end{aligned} \tag{4.54}$$

where $\Pr(\Delta N_{k,t} = i) = \exp(-\lambda_k \Delta t)(\lambda_k \Delta t)^i / i!$ for $k = 1, 2, 3$. For estimation purposes, the infinite sums need to be truncated and for the parameter estimates found here, the probability of more than 2 jumps per day is extremely small. As mentioned previously, to reduce the computational burden in estimation, the sums are truncated to 2 jumps per day. When jumps are truncated to two, the integrals in (4.52) and (4.53) can be evaluated to give the conditional densities of $\eta_{1,t}$ and $\eta_{2,t}$

[52] Although both of the short-term factors have mean-reversion factors, it is assumed that the jumps occurred sufficiently close to the current time step so that the mean-reversion only impacts jumps in the subsequent time-step.

as

$$p(\eta_{1,t}|\Delta N_{1,t}=1)=\omega_u\exp\Big(\frac{1}{2}\omega_u^2\sigma_1^2\Delta t-\omega_u\eta_{1,t}\Big)\Phi\Big(\frac{\eta_{1,t}-\omega_u\sigma_1^2\Delta t}{\sigma_1\sqrt{\Delta t}}\Big), \tag{4.55a}$$

$$p(\eta_{1,t}|\Delta N_{1,t}=2)=\omega_u^2\sigma_1\sqrt{\Delta t}\exp\Big(\frac{1}{2}\omega_u^2\sigma_1^2\Delta t-\eta_{1,t}\omega_u\Big)\Bigg[\phi\big(\eta_{1,t};\omega_u\sigma_1^2\Delta t,\sigma_1\sqrt{\Delta t}\big)\times\sigma_1\sqrt{\Delta t}+\frac{\eta_{1,t}-\omega_u\sigma_1^2\Delta t}{\sigma_1\sqrt{\Delta t}}\Phi\Big(\frac{\eta_{1,t}-\omega_u\sigma_1^2\Delta t}{\sigma_1\sqrt{\Delta t}}\Big)\Bigg], \tag{4.55b}$$

$$p(\eta_{2,t}|\Delta N_{2,t}=1)=\omega_d\exp\Big(\frac{1}{2}\omega_d^2\sigma_2^2\Delta t+\omega_d\eta_{2,t}\Big)\Phi\Big(\frac{-\eta_{2,t}-\omega_d\sigma_2^2\Delta t}{\sigma_2\sqrt{\Delta t}}\Big), \tag{4.55c}$$

$$p(\eta_{2,t}|\Delta N_{2,t}=2)=\omega_d^2\sigma_2\sqrt{\Delta t}\exp\Big(\frac{1}{2}\omega_d^2\sigma_2^2\Delta t+\eta_{2,t}\omega_d\Big)\Bigg[\phi\big(\eta_{2,t};-\omega_d\sigma_2^2\Delta t,\sigma_2\sqrt{\Delta t}\big)\times\sigma_2\sqrt{\Delta t}-\frac{\eta_{2,t}+\omega_d\sigma_2^2\Delta t}{\sigma_2\sqrt{\Delta t}}\Phi\Big(\frac{-\eta_{2,t}-\omega_d\sigma_2^2\Delta t}{\sigma_2\sqrt{\Delta t}}\Big)\Bigg], \tag{4.55d}$$

where $\Phi(x)$ denotes the standard normal cumulative distribution function.

For the densities of $\eta_t$ given above, the partial derivative of the log-density are,

$$\frac{\partial\log p(\eta_t)}{\partial\eta_t}=\frac{1}{p(\eta_t)}\frac{\partial p(\eta_t)}{\partial\eta_t}. \tag{4.56}$$

Assuming that the jumps are truncated at 2 for each observation, then the partial derivatives for each factor can be easily derived from (4.52) - (4.55) to give the following for $k=1,2,3$,

$$\frac{\partial p(\eta_{k,t})}{\partial\eta_{k,t}}=\Pr(\Delta N_{k,t}=0)\frac{\partial p(\eta_{k,t}|N_{k,t}=0)}{\partial\eta_{k,t}}+\Pr(\Delta N_{k,t}=1)\frac{\partial p(\eta_{k,t}|\Delta N_{k,t}=1)}{\partial\eta_{k,t}}+\Pr(\Delta N_{k,t}=2)\frac{\partial p(\eta_{k,t}|\Delta N_{k,t}=2)}{\partial\eta_{k,t}}, \tag{4.57}$$

where

$$\frac{\partial p(\eta_{1,t}|\Delta N_{1,t}=0)}{\partial\eta_{1,t}}=\frac{-\eta_{1,t}}{\sigma_1^2\Delta t}\phi(\eta_{1,t};0,\sigma_1^2\Delta t), \tag{4.58a}$$

$$\frac{\partial p(\eta_{1,t}|\Delta N_{1,t}=1)}{\partial \eta_{1,t}} = \omega_u \exp\Big(-\frac{1}{2}\omega_u^2\sigma_1^2\Delta t - \omega_u\eta_{1,t}\Big)\Bigg[\phi\big(\eta_{1,t};\omega_u\sigma_1^2\Delta t, \sigma_1\sqrt{\Delta t}\big) - \omega_u\Phi\Big(\frac{\eta_{1,t}-\omega_u\sigma_1^2\Delta t}{\sigma_1\sqrt{\Delta t}}\Big)\Bigg], \tag{4.58b}$$

$$\frac{\partial p(\eta_{1,t}|\Delta N_{1,t}=2)}{\partial \eta_{1,t}} = \omega_u^2\sigma_1\sqrt{\Delta t}\exp\Big(\frac{1}{2}\omega_u^2\sigma_1^2\Delta t - \eta_{1,t}\omega_u\Big) \times\Bigg\{-\omega_u\Bigg[\phi\big(\eta_{1,t};\omega_u\sigma_1^2\Delta t,\sigma_1\sqrt{\Delta t}\big)\sigma_1\sqrt{\Delta t} + \frac{\eta_{1,t}-\omega_u\sigma_1^2\Delta t}{\sigma_1\sqrt{\Delta t}} \times\Phi\Big(\frac{\eta_{1,t}-\omega_u\sigma_1^2\Delta t}{\sigma_1\sqrt{\Delta t}}\Big)\Bigg]-\frac{1}{\sigma_1\sqrt{\Delta t}}\Phi\Big(\frac{\eta_{1,t}-\omega_u\sigma_1^2\Delta t}{\sigma_1\sqrt{\Delta t}}\Big)\Bigg\}, \tag{4.58c}$$

$$\frac{\partial p(\eta_{2,t}|\Delta N_{2,t}=0)}{\partial \eta_{2,t}} = \frac{-\eta_{2,t}}{\sigma_2^2\Delta t}\phi(\eta_{2,t};0,\sigma_2^2\Delta t), \tag{4.58d}$$

$$\frac{\partial p(\eta_{2,t}|\Delta N_{2,t}=1)}{\partial \eta_{2,t}} = \omega_d \exp\Big(\frac{1}{2}\omega_d^2\sigma_2^2\Delta t + \omega_d\eta_{2,t}\Big)\Bigg[\phi\big(\eta_{2,t};-\omega_d\sigma_2^2\Delta t, \sigma_2\sqrt{\Delta t}\big) + \omega_d\Phi\Big(\frac{-\eta_{2,t}-\omega_d\sigma_2^2\Delta t}{\sigma_2\sqrt{\Delta t}}\Big)\Bigg], \tag{4.58e}$$

$$\frac{\partial p(\eta_{2,t}|\Delta N_{2,t}=2)}{\partial \eta_{2,t}} = \omega_d^2\sigma_2\sqrt{\Delta t}\exp\Big(\frac{1}{2}\omega_d^2\sigma_2^2\Delta t + \eta_{2,t}\omega_d\Big) \times\Bigg\{\omega_d\Bigg[\phi\big(\eta_{2,t};-\omega_d\sigma_2^2\Delta t,\sigma_2\sqrt{\Delta t}\big)\sigma_2\sqrt{\Delta t} - \frac{\eta_{2,t}+\omega_d\sigma_2^2\Delta t}{\sigma_2\sqrt{\Delta t}} \times\Phi\Big(\frac{-\eta_{2,t}-\omega_d\sigma_2^2\Delta t}{\sigma_2\sqrt{\Delta t}}\Big)\Bigg]-\frac{1}{\sigma_2\sqrt{\Delta t}}\Phi\Big(\frac{-\eta_{2,t}-\omega_d\sigma_2^2\Delta t}{\sigma_2\sqrt{\Delta t}}\Big)\Bigg\}, \tag{4.58f}$$

$$\frac{\partial p(\eta_{3,t}|\Delta N_{3,t}=n)}{\partial \eta_{3,t}} = \frac{-\eta_{3,t}}{\sigma_3^2\Delta t}\phi(\eta_{3,t};0,\sigma_3^2\Delta t + n\sigma_j^2). \tag{4.58g}$$

As $Q_t$ requires $\eta_t$, the approximating model can not be solved for analytically. Durbin and Koopman (1997) instead suggest that the approximating model can be found by using the following iterative procedure,

1. Generate a trial value for $\eta_t$, $t = 1, \ldots, T$.

2. Using (4.49) compute $Q_t$.

3. Using the Kalman filter and smoother on the linear state space model of the

approximating model, a new estimate of $\eta_t$ is derived.

4. Repeat steps 2 and 3 until convergence.

Having derived the approximating Gaussian model for importance sampling, we then need to calculate the likelihood function for parameter estimation. This is achieved using Monte Carlo simulations from the approximating Gaussian model. It is shown in Durbin and Koopman (1997) that for a given parameter set, $\Theta$, the likelihood can be expressed as,

$$L(\Theta) = L_G(\Theta)\mathbb{E}_G\left[\frac{p(x,y)}{q(x,y)}\right], \tag{4.59}$$

which for the jump-diffusion model reduces to

$$L(\Theta) = L_G(\Theta)\mathbb{E}_G\left[\frac{p(\eta)}{q(\eta)}\right]. \tag{4.60}$$

By sampling from the approximating Gaussian model, the expectation can be approximated by Monte Carlo simulation and so an estimate of the likelihood is given by

$$\hat{L}(\Theta) = L_G(\Theta)\frac{1}{T}\sum_{i=1}^{T}\frac{p(\eta^i)}{q(\eta^i)}, \tag{4.61}$$

where $\eta^i$ denotes a draw from $q(\eta)$ and the densities for each state variable, $p_(\eta_{i,t})$, $i = 1, 2, 3$, are given by (4.52) - (4.54). To obtain the parameter estmates for this model, the likelihood function or alternatively the log-likelihood function is maximized using the BFGS procedure. Each random draw used in the likelihood evaluation is obtained using a simulation smoother and in this case, the Durbin and Koopman (2002) method is used to sample from the approximating Gaussian density.

# Appendix 4.C The Unscented Kalman Filter

The unscented Kalman filter (UKF) relies on a scaled unscented transformation[53] for propagating the states through a nonlinear function. Consider the state space model given by (1.2) in Chapter 1, Section 1.1 where the function $h(\cdot)$ is nonlinear. For the application considered here, it is assumed that the measurement errors are additive, i.e.

$$y_t = h(x_t) + \epsilon_t,\ \epsilon_t \sim N(0, H_t). \tag{4.62}$$

For the state variable, it is assumed that $x_t$ has mean $\bar{x}_t$ and covariance $P_t$ although the errors are nonlinear with respect to the state variables. The scaled unscented transform estimates the first two moments of $y_t$ using a set of deterministically chosen points called "sigma" points. Given that the measurement errors are additive and the state errors are not, the state variable is redefined to be augmented with the state errors[54] so that $x_t^a = [x_t' \ \eta_t']'$ where $x_t^a \in \mathbb{R}^{m^a}$, $m^a = 2m$ and $\eta_t \sim N(0, Q_t)$. Therefore, the mean of $x_t^a$ is $\bar{x}_t^a = [\bar{x}_t' \ 0]'$ and its covariance is

$$P_t^a = \begin{bmatrix} P_t & 0 \\ 0 & Q_t \end{bmatrix}.$$

The scaled unscented transform is applied to the augmented state variable $x_t^a$ using $2m+1$ sigma points. The sigma points are chosen to match the mean and variance

[53] The original UKF relied on an unscented transform but could result in estimating a posterior covariance matrix that was non positve semi-definite. The scaled unscented transformation was introduced to address this issue.

[54] In the case of non-additive measurement errors, the state variable would also be augmented with the measurement errors. For further details on the more general UKF, readers are referred to Julier and Uhlmann (1997) and Wan and van der Merwe (2001).

of $x_t^a$ (accurate to the second order) in the following manner:

$$\mathcal{X}_{i,t} = \bar{x}_t^a, \tag{4.63a}$$

$$\mathcal{X}_{i,t} = \sqrt{(m^a + \lambda)P_{i,t}^a}\ ,\ i = 1, \ldots, m, \tag{4.63b}$$

$$\mathcal{X}_{i,t} = \sqrt{(m^a - \lambda)P_{i-m,t}^a}\ ,\ i = m+1, \ldots, 2m, \tag{4.63c}$$

for $t = 1, \ldots, T$ where $\sqrt{(m^a - \lambda)P_{i,t}^a}$ denotes the i'th column of the Cholesky factorisation of the matrix $(m^a - \lambda)P_t^a$. Each of the sigma points is associated with weights that depend on certain control parameters. These parameters are $\kappa$, $\alpha$, $\beta$ and have the following properties for calibrating the UKF: $\kappa \geq 0$ guarantees positive semi-definiteness of the covariance matrix, $0 \geq \alpha \geq 1$ controls the size of the sigma point distribution, and $\beta \geq 0$ gives partial control of higher order errors related to the estimation of the covariance matrix. For a Gaussian $x_t$, it is shown in Julier (2002) that $\beta = 2$ minimizes the higher order errors (although does not eliminate) when estimating the covariance matrix. Setting the additional parameter, $\lambda = \alpha^2(m^a + \kappa) - m^a$, the weights for the sigma points are

$$\mathcal{W}_0^m = \frac{\lambda}{m^a + \lambda}, \tag{4.64a}$$

$$\mathcal{W}_0^c = \frac{\lambda}{m^a + \lambda} + (1 - \alpha^2 + \beta), \tag{4.64b}$$

$$\mathcal{W}_i^m = \mathcal{W}_i^c = \frac{1}{2(m^a + \lambda)}\ ,\ i = 1, \ldots, 2m, \tag{4.64c}$$

where $\mathcal{W}_i^m$ and $\mathcal{W}_i^c$ denote the weights associated with calculating the mean and covariance respectively. Given the sigma points, their associated weights and the initial state, $x_0^a$ with mean $\bar{x}_0^a$ and covariance $P_0^a$, the UKF prediction step is given by,

$$\mathcal{X}_{i,t|t-1}^x = f(\mathcal{X}_{i,t-1}^x, \mathcal{X}_{i,t-1}^\eta), \tag{4.65}$$

$$\bar{x}_{t|t-1} = \sum_{i=0}^{2m} \mathcal{W}_i^m \mathcal{X}_{i,t|t-1}^x, \tag{4.66}$$

$$P_{t|t-1} = \sum_{i=0}^{2m} \mathcal{W}_i^c (\mathcal{X}_{i,t|t-1}^x - \bar{x}_{t|t-1})(\mathcal{X}_{i,t|t-1}^x - \bar{x}_{t|t-1})', \tag{4.67}$$

The update step is then,

$$\mathcal{Y}_{i,t|t-1} = h(\mathcal{X}_{i,t|t-1}^x), \tag{4.68}$$

$$\bar{y}_{t|t-1} = \sum_{i=0}^{2m} \mathcal{W}_i^m \mathcal{Y}_{t|t-1}, \tag{4.69}$$

$$v_t = y_t - \bar{y}_{t|t-1}, \tag{4.70}$$

$$F_t = \sum_{i=0}^{2m} \mathcal{W}_i^c (\mathcal{Y}_{i,t|t-1} - \bar{y}_{t|t-1})(\mathcal{Y}_{i,t|t-1} - \bar{y}_{t|t-1})' + H_t, \tag{4.71}$$

$$P_{x_t y_t} = \sum_{i=0}^{2m} \mathcal{W}_i^c (\mathcal{X}_{i,t|t-1}^x - \bar{x}_{t|t-1})(\mathcal{Y}_{i,t|t-1} - \bar{y}_{t|t-1})', \tag{4.72}$$

$$K_t = P_{x_t y_t} F_t^{-1}, \tag{4.73}$$

$$\bar{x}_t = \bar{x}_{t|t-1} + K_t v_t, \tag{4.74}$$

$$P_t = P_{t|t-1} + K_t F_t K_t', \tag{4.75}$$

where $K_t$ is the Kalman gain matrix.

# Chapter 5

# Summary and Conclusion

This thesis analysed approaches to estimate stochastic volatility models and investigated modelling techniques for pricing derivatives in commodity markets. The studies undertaken demonstrated application techniques, which have been traditionally applied to financial markets, to commodity markets. This included estimating stochastic volatility models, analysing risk premiums and pricing futures and options.

The first component of the thesis investigated discrete-time stochastic volatility modelling and compared three methods to estimate the models: a quasi-maximum likelihood estimator, a Monte Carlo likelihood estimator and a particle filter estimator. The analysis was conducted empirically using short-term interest rate and crude oil price data and considered simple extensions to the standard SV model. The analysis consisted of evaluating the in-sample and out-of-sample performance of the SV model using a number of diagnostic tests. This included an analysis of the parameter and conditional volatility estimates generated by the models and an investigation of the impact that the estimation procedure has on the estimates. Through the use of tests which evaluate the distributional properties of the model, the findings suggest that the choice of estimation procedure is more important than the choice of model when extracting the latent volatility process. For estimating SV models, the MCL procedure was found to perform the best. Also, based on

diagnostic tests in-sample and out-of-sample, the traditional AR(1) formulation of the SV model was found to be sufficient for modelling the volatility of short-term interest rates and crude oil prices.

The second component investigated a three-factor short/long factor model for commodity prices with the assumption of a time-varying market price of risk specification. The model was estimated using crude oil futures price data traded on NYMEX and demonstrated its performance over a two-factor model. The study found evidence to suggest time-varying risk premiums exist in the crude oil futures market and that they can help to improve forecasting performance. After extracting a time-series of the risk premium, the study also investigated whether risk premiums can be explained by risk factors in equity and bond markets. The evidence suggests that a portion of the risk premium in the crude oil futures market is correlated with equity and bond markets.

The third component extended the short/long factor model with jumps and stochastic volatility. Semi-closed form solutions to futures and European option prices were derived for the model using the properties of affine jump-diffusion models. The model was empirically evaluated using daily futures and options data traded on NYMEX. An initial study investigated whether options prices can be inferred from futures prices. Although it was found that adding jumps or stochastic volatility explains the distributional properties of crude oil prices better, this did not translate into significantly better option pricing performance. Thus, estimating the models using both futures and options prices, it was found that stochastic volatility is necessary for fitting option prices over the maturity and moneyness dimensions. However, whilst jumps without the presence of stochastic volatility did not improve options pricing much, it was found necessary in the SVJ model to fit the steep implied volatility smiles exhibited by short term options.

The implications of this thesis apply to the areas of financial modelling, risk management and portfolio investment. The first component demonstrated methods for evaluating and estimating risk using SV models in fixed income and commodity markets. It analysed different methods to estimate volatility and volatility models, as well as determine the accuracy of volatility forecasts. The second component considered an intuitive modelling approach for commodities and found evidence of time-varying risk premiums in the crude oil futures market. An analysis of the risk premiums demonstrates its potential use in portfolio diversification with other financial assets. The third component proposed a model for futures and option pricing which can recover market implied volatilities. By providing realistic behaviour of both spot and derivative prices, the model can be used in risk management applications such as determining hedge ratios for commodity option portfolios.

There are a number of areas in which the research can be expanded. Firstly, the empirical analysis has focussed mainly on the crude oil futures market. The empirical analysis in each of the studies can be easily expanded to other commodities, especially other energy commodities such as other petroleum products, heating oil and natural gas. Also, in terms of the estimation of the commodity risk premiums, only price risk premiums were investigated. With the stochastic volatility jump-diffusion model introduced in Chapter 4, both volatility risk and jump risk can be extracted and analysed.

Furthermore, the SVJ model can be extended to incorporate features considered in the literature in other financial markets. As the model falls under the affine jump-diffusion class of models, this gives scope for the model to be extended within this class. For instance, the jumps in the model were assumed to have a constant intensity. However, the jump intensity can be allowed to vary over time which could possibly allow better fits to option prices over maturity and moneyness. This also

allows for fitting of time-varying skew in option prices which has been exhibited in other markets. Lastly, as mentioned in Chapter 4, there is evidence for unspanned stochastic volatility in the crude oil futures market. The SVJ model could be reformulated so that it accommodates unspanned stochastic volatility. Ideally, a model would be able to account for all of these dynamics. However, as mentioned in the introduction, the complexities involved with these features increase the difficulty of both pricing derivatives and calibrating the models to market data. This motivates further research in these areas.

# Bibliography

Aït-Sahalia, Y., 2002. Maximum Likelihood Estimation of Discretely Sampled Diffusions: A Closed-form Approximation Approach. *Econometrica*, 70(1), 223-262.

Aiube, F. A. L., Baidya, T. K. N., Tito, E. A. H., 2008. Analysis of Commodity Prices with the Particle Filter. *Energy Economics*, 20, 597-605.

Alquist, R., Kilian, L., 2010. What do we Learn from the Price of Crude Oil Futures? *Journal of Applied Econometrics*, 25, 539-573.

Andersen, T. G., Davis, R. A., Kreiss, J. P., Mikosch, T., 2009. *Handbook of Financial Time Series*. Springer, New York.

Askari, H., Krichene, N., 2008. Oil Price Dynamics (2002-2006). *Energy Economics*, 30, 2134-2153.

Bailey, W., Chan, K., 1993. Macroeconomic Influences and the Variability of the Commodity Futures Basis. *Journal of Finance*, 48, 555-573.

Bakshi, G., Carr, P., Wu, L., 2008. Stochastic Risk Premiums, Stochastic Skewness in Currency Options and Stochastic Discount Factors in International Economies. *Journal of Financial Economics*, 87, 132-56.

Bakshi, G., Madan, D., 2000. Spanning and Derivative-Security Valuation. *Journal of Financial Economics*, 55, 205-238.

Ball, C. A., Torous, W. N., 1999. The Stochastic Volatility of Short-Term Interest Rates: Some International Evidence. *Journal of Finance*, 54(6), 2339-2359.

Barone-Adesi, G., Whaley, R., 1987. Efficient Analytic Approximation of American Option Values. *Journal of Finance*, 42, 301-20.

Bates, D. S., 1996. Jumps and Stochastic Volatility: Exchange Rate Processes Implicit in Deusche Mark Options. *Review of Financial Studies*, 9, 69-107.

Baum, C. F., Barkoulas, J., 1996. Time-varying Risk Premia in the Foreign Currency Futures Basis. *Journal of Futures Markets*, 16, 735-755.

Berkowitz J., 2001. Testing Density Forecasts with Applications to Risk Management. *Journal of Business and Economic Statistics*, 19, 465-474.

Bessembinder, H., and Kalok C., 1992. Time-varying Risk Premia and Forecastable Returns in Futures Markets. *Journal of Financial Economics*, 32, 169-193.

Bibkov, R., Chernov, M., 2009. Unspanned Stochastic Volatility in Affine Models: Evidence from Eurodollar Futures and Options. *Management Science*, 55, 1292-1305.

Black, F., 1976. The Pricing of Commodity Contracts. *Journal of Financial Economics*, 3, 167-180.

Black, F., Scholes, M., 1973. The Pricing of Options and Corporate Liabilities. *Journal of Political Economy*, 81, 637654.

Bollerslev, T., 1986. Generalized Autoregressive Conditional Heteroscedasticity. *Journal of Econometrics*, 31, 307-327.

Brennan, M. J., 1958. The Supply of Storage. *American Economic Review*, 47, 50-72.

Broadie, M., Chernov, M., Johannes, M., 2007. Model Specification and Risk Premiums: Evidence from Futures Options. *Journal of Finance*, 62, 1453-1490.

Broto, C., Ruiz, E., 2004. Estimation Methods for Stochastic Volatility Models: A Survey. *Journal of Economic Surveys*, 18, 613-649.

Cappé, O., Godsill, S. J., Moulines, E., 2007. An Overview of Existing Methods and Recent Advances in Sequential Monte Carlo. *Proceedings of the IEEE*, 95(5), 899-924.

Carr, P., Wu, L., 2007. Stochastic Skew in Currency Options. *Journal of Financial Economics*, 86, 213-247.

Cartea, Á., Figueroa, M. G., 2005. Pricing in Electricity Markets: A Mean Reverting Jump Diffusion Model with Seasonality. *Applied Mathematical Finance*, 12, 313-335.

Cartea, Á., Williams, T., 2008. UK Gas Markets: The Market Price of Risk and Applications to Multiple Interruptible Supply Contracts. *Energy economics*, 30, 829-846.

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

Chacko, G., Das, S., 2002. Pricing Interest Rate Derivatives: A General Approach. *Review of Financial Studies*, 15(1), 195-241.

Chan, K.C., Karolyi, G. A., Longstaff, F., Sanders, A., 1992. The Volatility of Short-Term Interest Rates: An Empirical Comparison of Alternative Models of Term Structure of Interest Rates. *Journal of Finance*, 47, 1209-1227.

Chen, R., Scott, L., 2003. Multi-factor Cox-Ingersoll-Ross Models of the Term Structure: Estimates and Tests from a Kalman Filter Model. *Journal of Real Estate Finance Economics*, 27, 143-172.

Cheridito, P., Filipović, D., Kimmel, R. L., 2007. Market Price of Risk Specifications for Affine Models: Theory and Evidence. *Journal of Financial Economics*, 83, 123-170.

Cheridito, P., Filipović, D., Yor, M., 2005. Equivalent and Absolutely Continuous Measure Changes for Jump-Diffusion Processes. *Annals of Applied Probability*, 15(3), 1713-1732.

Considine, T. J., Larson, D. F., 2001. Risk Premiums on Inventory Assets: The Case of Crude Oil and Natural Gas. *Journal of Futures Markets*, 21, 109-126.

Cont, R., Tankov, P., 2004. *Financial Modelling with Jump Processes*. Chapman and Hall / CRC Financial mathematics series.

Cortazar, G., Naranjo, L., (2006). An N-factor Gaussian Model of Oil Futures Prices. *Journal of Futures Markets*, 26, 243-268.

Cox, J. C., Ingersoll, J. E., Ross, S. A., 1981. The Relation Between Forward Prices and Futures Prices. *Journal of Financial Economics*, 9, 32146.

Cox, J. C., Ingersoll, J. E., Ross, S. A., 1985. A Theory of the Term Structure of Interest Rates. *Econometrica*, 53, 385-407.

Crosby, J., 2008. A Multi-Factor Jump-Diffusion Model for Commodities. *Quantitative Finance*, 8(2), 181-200.

Dai, Q., Singleton, K., 2000. Specification Analysis of Affine Term Structure Models. *Journal of Finance*, 55(5), 1943-1978.

Danielsson, J., 1994. Stochastic Volatility in Asset Prices: Estimation with Simulated Maximum Likelihood. *Journal of Econometrics*, 64, 375-400.

Das, S.R., 2002. The Surprise Element: Jumps in Interest Rates. *Journal of Econometrics*, 106, 27-65.

de Jong, P., Shephard, N., 1995. The Simulation Smoother for Time Series Models. *Biometrika*, 82, 339-350.

Deaton, A., Laroque, G., 1996. Competitive Storage and Commodity Price Dynamics. *Journal of Political Economy*, 104, 896-923.

Deaves, R., Krinsky, I., 1992. Risk Premiums and Efficiency in the Market for Crude Oil Futures. *Energy Journal*, 13, 93-117.

Dempster, M. A. H., Medova, E., Tang, K., 2008. Long and Short Term Jumps in Commodity Futures Prices. Working paper, available at SSRN: http://ssrn.com/abstract=1107966.

Doran, J.S., Ronn, E.I., 2008. Computing the Market Price of Volatility Risk in the Energy Commodity Markets. *Journal of Banking and Finance*, 32, 2541-2552.

Doucet, A., de Freitas, J. F. G., Gordon, N., 2001. *Sequential Monte Carlo Methods in Practice*, Springer-Verlag.

Doucet, A., Godsill, S., Andrieu, C., 2000. On Sequential Monte Carlo Sampling Methods for Bayesian Filtering. *Statistics and Computing*, 10, 197-208.

Duan, J. C., Simonato, J. G., 1999. Estimating and Testing Exponential-Affine Term Structure Models by Kalman Filter. *Review of Quantitative Finance and Accounting*, 13, 111-135.

Duffee, G.R., 2002. Term premia and Interest Rate Forecasts in Affine Models. *Journal of Finance*, 57, 405-443.

Duffie, D., Pan, J., Singleton, K., 2000. Transform Analysis and Asset Pricing for Affine Jump-Diffusions. *Econometrica*, 68 (6), 1343-1376.

Durbin, J., Koopman, S. J., 1997. Monte Carlo Maximum Likelihood Estimation of non-Gaussian State Space Model. *Biometrika*, 84(3), 669-684.

Durbin, J., Koopman, S. J., 2001. *Time Series Analysis by State Space Methods*, Oxford University Press.

Durbin, J., Koopman, S. J., 2002. A Simple and Efficient Simulation Smoother for State Space Time Series Analysis. *Biometrika*, 89(3), 603-615.

Fama, E. F., French, K. R., 1987. Commodity Futures Prices: Some Evidence on Forecast Power, Premiums, and the Theory of Storage. *Journal of Business*, 60, 55-73.

Fama, E. F., French, K. R., 1989. Business Conditions and Expected Returns on Stocks and Bonds. *Journal of Financial Economics*, 25, 23-50.

Geman, H., 2005. *Commodities and Commodity Derivatives: Modeling and Pricing for Agriculturals, Metals and Energy*. John Wiley & Sons.

Gibson, R., Schwartz, E., 1990. Stochastic Convenience Yield and the Pricing of Oil Contingent Claims. *Journal of Finance*, 45, 959-976.

Godambe, V. P., 1985. The Foundations of Finite Sample Estimation in Stochastic Processes. *Biometrika*, 72(2), 419-428.

Gordon, N., Salmond, D., Smith, A. F., 1993. Novel Approach to Nonlinear/non-Gaussian Bayesian State Estimation. *IEE Proceedings F, Radar Signal Processes*, 140, 107-113.

Harrison, J.M., Pliska, S.R., 1981. Martingales and Stochastic Integrals in the Theory of Continuous Trading. *Stochastic Processes and their Applications*, 11, 215-260.

Harvey, A.C., 1989. *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge University Press.

Harvey, A.C., Ruiz, E., Shephard, N., 1994. Multivariate Stochastic Variance models. *Review of Economic Studies*, 61, 247-264.

Harvey, A.C., Shephard, N., 1993. The Econometrics of Stochastic Volatility. London School of Economics, Working Paper.

Heath, D., Jarrow, R., Morton, A., 1992. Bond Pricing and the Term Structure of Interest Rates: A New Methodology. *Econometrica*, 60, 77-105.

Heston, S. L., 1993. A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options. *The Review of Financial Studies*, 6(2), 327-343.

Hikspoors, S., Jaimungal, S., 2008. Asymptotic Pricing of Commodity Derivatives using Stochastic Volatility Spot Models. *Applied Mathematical Finance*, 15(56), 449-477.

Hughen, W.K., 2010. A Maximal Affine Stochastic Volatility Model of Oil Prices. *The Journal of Futures Markets*, 30, 101-133.

Jacquier, E., Polson, N.G., Rossi, P.E., 1994. Bayesian Analysis of Stochastic Volatility Models. *Journal of Business and Economic Statistics*, 12, 371-417.

Johannes, M., Polson, N., Stroud, J., 2009. Optimal Filtering of Jump-Diffusions: Extracting latent States from Asset Prices. *The Review of Financial Studies*, 22(7), 2759-2799.

Jones, C.M., Lamont, O., Lumsdaine, R., 1998. Macroeconomic News and Bond Market Volatility. *Journal of Financial Economics*, 47, 315-337.

Julier, S. J., Uhlmann, J. K., 1997. A New Extension to the Kalman Filter to Nonlinear Systems. *In Proceedings of AeroSense: 11th International Symposium Aerospace/Defense Sensing, Simulation and Controls*, 182-193.

Julier, S. J., 2002. The Scaled Unscented Transformation. *In Proceedings of the American Control Conference*, 4555-4559.

Julier, S. J., Uhlmann, J. K., 2004. Unscented Filtering and Nonlinear Estimation. *Proceedings of the IEEE*, 92(3), 401-422.

Kaldor, N., 1939. Speculation and Economic Stability. *Review of Economic Studies*, 7, 1-27.

Kalimipali, M., Susmel, R., 2004. Regime-Switching Stochastic Volatility and Short-Term Interest Rates. *Journal of Empirical Finance*, 11, 309-329.

Kalman, R. E., 1960, A New Approach to Linear Filtering and Prediction Problems. *Transactions of the ASMEJournal of Basic Engineering*, 82 (Series D), 35-45.

Keim, D., Stambaugh, R., 1986. Predicting Returns in the Stock and Bond Markets. *Journal of Financial Economics*, 17,357-390

Keynes, J.M., 1930. *The Applied Theory of Money*. Macmillan & Co., London.

Kim, S., Shephard, N., Chib, S., 1998. Stochastic Volatility: Likelihood Inference and Comparison with ARCH Models. *Review of Economic Studies*, 65, 361-393.

Kitagawa, G., 1996. Monte Carlo Filter and Smoother for Non-Gaussian, Nonlinear State Space Models. *Journal of Computational and Graphical Statistics*, 5(1), 1-25.

Kolos, S.P., Ronn, E.I., 2008. Estimating the Commodity Market Price of Risk for Energy Prices. *Energy Economics*, 30, 621-641.

Koopman, S.J., Lee, K.M., 2004. Estimating Stochastic Volatility Models: A Comparison of Two Importance Samplers. *Studies in Nonlinear Dynamics & Econometrics*, 8(2), 5

Kou, S. G., 2002. A Jump-Diffusion Model for Option Pricing. *Management Science*, 48 (8), 1086-1101.

Lamoureux, C.G., Paseka, A., 2009. Information in Option Prices and the Underlying Asset Dynamics. Working Paper.

Liesenfeld, R., Richard, J. F., 2003. Univariate and Multivariate Stochastic Volatility Models: Estimation and Diagnostics. *Journal of Empirical Finance*, 10, 505-531.

Melino, A., Turnbull, S. M., 1990. Pricing Foreign Currency Options with Stochastic Volatility. *Journal of Econometrics*, 45, 239-265.

Merton, R., 1973. Theory of Rational Option Pricing. *The Bell Journal of Economics and Management Science*, 4(1), 141-183.

Merton, R., 1976. Option Pricing when Underlying Stock Returns are Discontinuous. *Journal of Financial Econometrics*, 3, 125-144.

Moosa, I. A., Al-Loughani, N. E., 1994. Unbiasedness and Time Varying Risk Premia in the Crude Oil Futures Market. *Energy Economics*, 16, 99-105.

Nelson, C., Siegel, A., 1987. Parsimonious Modeling of Yield Curves. *Journal of Business*, 60, 473-489.

Newey, W. K., West, K. D., 1987. A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. *Econometrica*, 55, 703-708.

Pandher, G. S., 2001. Estimation of Excess Returns from Derivative Prices and Testing for Risk Neutral Pricing. *Econometric Theory*, 17, 785-819.

Pitt, M. K., 2002. Smooth Particle Filters for Likelihood Evaluation and Maximisation. The Warwick Economics Research Paper Series (TWERPS) 651, University of Warwick, Department of Economics.

Pitt M. K., Shephard, N., 1999. Filtering via Simulation: Auxiliary Particle Filters. *Journal of the American Statistical Association*, 94(446), 590-599.

Ramezani, C. A., Zend, Y., 2007. Maximum Likelihood Estimation of the Double Exponential Jump-Diffusion Process. *Annals of Finance*, 3, 487-507.

Richter, M. C., Sørensen, C., 2002. Stochastic Volatility and Seasonality in Commodity Futures and Options: The Case of Soybeans. Working Paper, Copenhagen Business School.

Rohatgi, V. K., 1976. *An Introduction to Probability Theory and Mathematical Statistics*. John Wiley and Sons Inc., New York.

Rosenblatt, M., 1952. Remarks on a Multivariate Transformation. *The Annals of Mathematical Statistics*, 23, 470-472.

Routledge, B. R., Seppi, D. J., Spatt, C. S., 2000. Equilibrium Forward Curves for Commodities. *Journal of Finance*, 55, 1297-1338.

Ruiz, E., 1994. Quasi-Maximum Likelihood Estimation of Stochastic Volatility Models. *Journal of Econometrics*, 63, 289-306.

Sadorsky, P., 2002. Time-varying Risk Premiums in Petroleum Futures Prices. *Energy Economics*, 24, 539-556.

Samuelson, P. A., 1965. Proof that Properly Anticipated Prices Fluctuate Randomly. *Industrial Management Review*, 6, 41-49.

Sandmann, G., Koopman, S.J., 1998. Estimation of Stochastic Volatility Models via Monte Carlo Maximum Likelihood. *Journal of Econometrics*, 87, 271-301.

Sapp, T. R. A., 2009. Estimating Continuous-Time Stochastic Volatility Models of the Short-Term Interest Rate: A Comparison of the Generalized Method of Moments and the Kalman Filter. *Journal of Quantitative Finance and Accounting*, 33(3), 303-326.

Schwartz, E. S., (1997), The Stochastic Behavior of Commodity Prices: Implications for valuation and hedging. *Journal of Finance*, 52, 923-973.

Schwartz, E. S., Smith, J. E., 2000. Short-Term Variations and Long-Term Dynamics in Commodity Prices. *Management Science*, 46, 893-911.

Smith, D. R., 2002. Markov Switching and Stochastic Volatility Diffusion Models of Short-Term Interest Rates. *Journal of Business and Economic Statistics*, 20(2), 183-197.

Sørenson, C., 2002. Modeling Seasonality in Agricultural Commodity Futures. *Journal of Futures Markets*, 22, 393-426.

Sun, L. C., 2005. Regime Shifts in Interest Rate Volatility. *Journal of Empirical Finance*, 12, 418-434.

Taylor, S. J., 1986. *Modelling Financial Time Series*. Chichester, UK: John Wiley and Sons.

Thavaneswaran, A., Abraham, B., 1988. Estimation for Non-Linear Time-Series Models using Estimating Equations. *Journal of Time Series Analysis*, 9(1), 99-108.

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.

Villaplana, P., 2002. Pricing Power Derivatives: A Two-Factor Jump-Diffusion Approach. Working Paper.

Vo, M. T., 2009. Regime-Switching Stochastic Volatility: Evidence from the Crude Oil Market. *Energy Economics*, 31(5),779-788.

Vuong, Q. H., 1989. Likelihood Ratio Tests for Model Selection and Non-Nested Hypotheses. *Econometrica*, 57, 307-333.

Wan, E. A., van der Merwe, R., 2001. The Unscented Kalman Filter. In: Haykin, S. (Ed.), *Kalman Filtering and Neural Networks*. Wiley, New York.

Weron, R., (2008). Market Price of Risk Implied by Asian-style Electricity Options and Futures. *Energy Economics*, 30, 1098-1115.

White, H., 1982. Maximum Likelihood Estimation of Misspecified Models. *Econometrica*, 50(1), 1-25.

Wong, B., Heyde, C. C., 2006. On Changes of Measure in Stochastic Volatility Models. *Journal of Applied Mathematics and Stochastic Analysis*, 1-13.

Working, H., 1949. The Theory of Price and Storage. *American Economic Review*, 39, 1254-1262.

Yan, X., 2002. Valuation of Commodity Derivatives in a New Multi-Factor Model. *Review of Derivatives Research*, 5, 251-271.