Oil prices – Brownian motion or mean reversion? A study using a one year ahead density forecast criterion.

Nigel Meade

Imperial College Business School

Published in Energy Economics 32 (2010) 1485–1498

Abstract. For oil related investment appraisal, an accurate description of the evolving uncertainty in the oil price is essential. For example, when using real option theory to value an investment, a density function for the future price of oil is central to the option valuation. The literature on oil pricing offers two views. The arbitrage pricing theory literature for oil suggests geometric Brownian motion and mean reversion models. Empirically driven literature suggests ARMA-GARCH models. In addition to reflecting the volatility of the market, the density function of future prices should also incorporate the uncertainty due to price jumps, a common occurrence in the oil market.

In this study, the accuracy of density forecasts for up to a year ahead is the major criterion for a comparison of a range of models of oil price behaviour, both those proposed in the literature and following from data analysis. The Kullbach Leibler information criterion is used to measure the accuracy of density forecasts.

Using two crude oil price series, Brent and West Texas Intermediate (WTI) representing the US market, we demonstrate that accurate density forecasts are achievable for up to nearly two years ahead using a mixture of two Gaussians innovation process with GARCH and no mean reversion.

Keywords. Time series, Density forecasting, commodity prices

J.E.L classification. C22, C53, G10.

Imperial College Business School, Exhibition Road, London SW7 2AZ
Tel +44 (0) 20 7594 9116; Fax +44 (0) 20 7594 7442
n.meade@imperial.ac.uk

Acknowledgement: The author would like to express his gratitude to three anonymous referees for their insightful comments on an earlier version of the paper which led to an improvement in its content and exposition.

1 Introduction

For investment appraisal where the price of oil is a source of uncertainty, a density function for the future price of oil is necessary. For example, if one were using real option theory to value an investment, an accurate description of the evolving uncertainty in the oil price is central to the option valuation. The need to provide this density function provides the motivation for this study. A secondary motivation is the increasing use of oil based derivatives, such as futures, for risk management in the oil market and for investment purposes by other investors. We will compare the models of oil price behaviour proposed in the literature, as well as additional models based on data analysis, using the accuracy of density forecasts for a year ahead as our main criterion. We choose this criterion because the accuracy of the density forecast from a given model is far more informative than the accuracy of a point forecast or of a prediction interval. A one year horizon is chosen as a convenient filter to identify potentially viable models. The accuracy of density forecasts from the surviving models is evaluated over longer horizons in order to find the maximum horizon for which the density is plausible.

A pioneering study of the empirical behaviour of commodity prices is Labys and Granger (1971). Theoretical models of commodity price behaviour are discussed and developed by Deaton and Laroque (1992). Looking specifically at oil price behaviour, the literature has two main themes for the choice of model. One modelling philosophy is dictated by arbitrage pricing theory where the motivation is to provide a pricing framework for a derivative or futures, see for example Brennan and Schwartz (1985) and Schwartz (1997). These continuous time models include geometric Brownian motion and mean reversion models. An alternative philosophy is data driven where a model is chosen from a universe of models such as the ARIMA framework according to a goodness of fit procedure. In an example of density forecasting, Melick and Thomas (1997) use option prices to recover the implied probability density of crude prices, since derivative prices encapsulate current uncertainty.

In order to successfully model the price of oil, the density function of future prices must be able to incorporate the uncertainty due to shocks as well as the underlying volatility of a stable market. The price of oil is subject to shocks caused by threats to stability in the Middle East, Africa, South America and elsewhere; by reports of lower than usual crude stocks in the US and by many other actual or anticipated events. The financial press contains explanations of oil price fluctuations on virtually a daily basis (a search for ‘oil price’ on the Financial Times website, ft.com, reveals over 38,000 articles between January 2003 and September 2008). We take the view that it is fruitless to model oil prices by considering specific events in detail for two reasons, firstly there are so many, secondly, modelling these events tends only to improve in sample fit rather than improve forecasting accuracy. Clements and Hendry (1996) identify strategies that are more likely to forecast better in the presence of structural breaks, the models we use follow in the spirit of their comments.

We use two oil price series, Brent representing the European oil market and West Texas Intermediate (WTI) representing the US market. Oil price data are easily available from 1946 onwards, however between 1946 and 1970 the oil price crept up from 1.63 US$ per barrel ($/bbl) to 3.39$/bbl. The more eventful history from 1970 onwards is summarised in Figure 1. Various Middle Eastern crises led to the price exceeding 10$/bbl in 1974. After reaching a peak of 37.50$/bbl in March 1981, the price fell back to the 10 - 20$/bbl band where it stayed until 2000. Stevens (2005) attributes the sharp price drop in 1986 to excess upstream capacity. In the late 1980s, the international oil companies (IOCs) began to use long term contracts less and to use the market prices of marker crude oils, such as Brent, WTI and Arab Light, as a basis to price other crude oils. The period 2002 onwards shows an upward movement in prices, during this period surplus production capacity decreases from 7 million bbl per day (Jan 2002) to less than one million bbl per day (Oct 2004). The 'oil price bubble' in mid 2008, (when the price of Brent and WTI both exceeded US$140, has been explained by some or all of the following causes: an economic boom in the world's largest developing countries, particularly China and India; restrictions in supply; known reserves are diminishing; an excess of speculative activity.[1]

The structure of the paper is as follows. We review pricing models in Section 2, where we also discuss the estimation of these models and review the density forecasting literature. In Section 3, we describe the data used and perform some exploratory analysis of the data. Model estimation and a comparison of density forecasting accuracy are performed in Section 4. We offer our conclusions in Section 5.

2 Review of Pricing Models

Theoretical models of commodity price behaviour are reviewed in Deaton and Laroque (1992). They develop the rational expectations competitive storage model of price formation for commodities and show that the commodity price belongs to one of two regimes; one where demand is equal to current supply and inventory and one where demand exceeds current supply. In the long run, the price oscillates between these regimes. Their empirical work used data for harvestable crop commodities plus some metals; neither crude oil nor oil products were included. In the short term, inventory levels have an effect on crude oil prices, see Pindyck (2004). The use of inventory data has been used for short term (up to three months ahead) forecasting by Ye, Zyren and Shore (2006). In the following two subsections, inventory data is not considered, this is because it is of lower frequency than the price series modelled. In addition, inventories are measured nationally, whereas the prices are for an international commodity.

2.1 Arbitrage Pricing Models

Brennan and Schwartz (1985) proposed an early oil pricing model as a component of an exercise to price a natural resource investment, in effect pricing a real option. They used a geometric Brownian motion model for the price S:

[1] M.Desai, Act now to prick the oil price bubble, Financial Times, 5 June 2008.

$$\frac{dS}{S} = \mu \, dt + \sigma \, dz_t \qquad (1)$$

where dz_t is the increment in a Gauss-Wiener process with drift μ and instantaneous standard deviation σ.

Schwartz (1997) proposed three pricing models with one, two and three factors. The factors are spot price, convenience yield and interest rate. The purpose of these models was to price futures, whereas here our pre-occupation is with the spot price (although this is the price of the futures contract closest to maturity). Thus we only need to consider the first of the three models proposed as the last two models collapse to (2) if we assume a constant convenience yield. The first model is a mean reversion model:

$$\frac{dS}{S} = \kappa\left(\eta - \ln\left(S\right)\right) dt + \sigma \, dz_t \qquad (2)$$

where the log of the oil price reverts to a long term mean η at a rate defined by κ.

Cortazar and Schwartz (1997, 1998) use these models in a real option evaluation of an undeveloped oil field. Smith and McCardle (1999) use a similar approach to evaluate oil and gas investments.

Schwartz and Smith (2000) propose a combined long term and short term model; the log price is the sum of these components:

$$\ln\left(S_t\right) = \chi_t + \xi_t \qquad (3)$$

The short term process is reversion process to a zero mean:

$$d\chi_t = -\kappa\chi_t dt + \sigma_\chi \, dz_\chi \qquad (4)$$

and the long term process is a geometric Brownian motion (in price):

$$d\xi_t = \mu_\xi dt + \sigma_\xi \, dz_\xi \qquad (5).$$

The two processes are correlated as follows

$$dz_\chi \, dz_\xi = \rho_{\chi\xi} dt \qquad (6).$$

The models described in (1), (2) and (3) to (6) are the three continuous time models to be found in the literature (to the best of our knowledge). Although other authors, for example Cortazar and Naranjo (2006) have proposed multi-factor models designed to capture the term structure of future contracts, they have not introduced another spot price process and it is this process that is our prime focus.

Typically derivative pricing uses a risk-neutral pricing measure where the asset price at time *t*, discounted at the risk-free rate *r*, $X_t = S_t \exp\left(-rt\right)$, is a martingale, (i.e. $E\left(X_{t+k} \mid \Theta_t\right) = X_t$, where $k > 0$). The martingale property ensures arbitrage-free pricing and that the market is complete, i.e. a contingent claim can be hedged by a portfolio of the asset and a risk-less bond. If the observed price process can be modelled using geometric Brownian motion, as above, then the conversion of this process to a martingale is relatively straightforward. The consequence is that pricing formulas for some contingent claims can be obtained in a closed form.

The Euler-Maruyama method (see Higham, 2001) is a well known procedure to convert continuous models to a discrete time formulation. If $Z_t = \ln\left(\frac{S_t}{S_{t-1}}\right)$, the Brownian motion of (2) becomes:

$$Z_t = \mu + \varepsilon_t \tag{7}$$

where $\varepsilon_t \sim N\left(0, \sigma^2\right)$. The mean reversion model in (3) becomes:

$$Z_t = \kappa\left(\eta - \ln\left(S_{t-1}\right)\right) + \varepsilon_t \qquad (8).$$

The short and long term processes in (3) to (6), simplify to (9) in discrete time if we assume a single source of error model:

$$Z_t = \kappa\left(\eta - \ln\left(S_{t-1}\right)\right) + \varepsilon_t + \theta\varepsilon_{t-1} \qquad (9).$$

2.2 Discrete Time Models

There have been several studies of crude oil or oil product price series using discrete time models. Kumar (1992) compared the forecasting accuracy of a random walk model and an ARMA model of oil prices with oil futures up to nine months ahead. He found that the futures were unbiased and more accurate than the ARMA model which was little different from the random walk. However, the accuracy of the futures deteriorated with horizon and was only marginally more accurate than the random walk for all horizons. Panas and Nini (2000) model oil product prices, their most general model is an AR(1) - GARCH(1,1) in mean model. That is:

$$Z_t = \mu + \phi_1 Z_{t-1} + \gamma\sigma_{t-1}^{\lambda} + \varepsilon_t \tag{10}$$

and the GARCH (1,1) process is

$$\sigma_t^2 = \alpha_0 + \alpha_1\varepsilon_t^2 + \beta_1\sigma_{t-1}^2 \tag{11}$$

where σ_t^2 is the conditional variance of Z_t, $\alpha_0 > 0, \alpha_1 > 0, \beta_1 \geq 0$ and λ is typically 1 or 2. In their study of fourteen products on the Rotterdam and Mediterranean markets, they found six series with significant autoregressive terms ($\phi_1 \neq 0$) and another six series displayed a significant GARCH in mean effect. ($\gamma \neq 0$). The GARCH effect was significant in all fourteen series. Cabedo and Moya (2003) used ARMA models to predict value at risk (VaR) for Brent crude. They found that that a ARMA(1,1) model gave better out of sample VaR results than the AR(1) – GARCH(1,1) model. Hung et al. (2007) used GARCH(1,1) with three different innovation processes for one day ahead forecasts of Brent and WTI crudes and three products, they found that a heavy tailed distribution most accurately captured this short term risk.

2.3 Extensions to pricing models to be used in this study

In this section, we consider the following general model which includes the important properties of the models discussed in the previous sections

$$Z_t = \mu + \kappa\left(\eta - \ln\left(S_{t-1}\right)\right) + \sum_{k=1}^{p} \phi_k Z_{t-1} + \varepsilon_t \quad (12).$$

where ε_t follows a possibly asymmetric GARCH (1,1) process

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_t^2 + \xi \min\left(\varepsilon_t, 0\right)^2 + \beta_1 \sigma_{t-1}^2 \quad (13)$$

σ_t^2 is the conditional variance of Z_t, $\alpha_0 > 0, \alpha_1 > 0, \beta_1 \geq 0$ and $\kappa\mu = 0$. The discrete version of Brownian motion is obtained by constraining $\kappa, p, \alpha_1, \xi, \beta_1$ to be zero, see (7). Mean reversion is modelled by constraining μ to be zero, see (8). Autoregressive terms can be included as in (10), the moving average term in (9) is not included but it can be represented by a high order AR process. Asymmetry in the volatility process of equity return series has been detected by, for example, Verhoeven and McAleer (2004), this asymmetry reflects the different response of the market to negative and positive shocks. For simplicity in modelling the mean of the return process, we omitted the GARCH in mean effect used in (10).

It is feasible that over a prolonged period, the mean to which a mean reversion process reverts may change over time, thus, in a further extension; we consider reversion to a time dependent mean,

$$Z_t = \mu + \kappa\left(\eta_{t-1} - \ln\left(S_{t-1}\right)\right) + \sum_{k=1}^{p} \phi_k Z_{t-1} + \varepsilon_t \quad (14).$$

In this process, the time dependent mean follows a moving average process

$$\eta_t = \eta_{t-1} + \gamma_t\left(\ln\left(S_{t-1}\right) - \eta_{t-1}\right) \quad (15)$$

and γ_t is estimated simultaneously with the other model parameters.

2.4 Alternative density functions to the Gaussian

As mentioned in the introduction, oil prices are often subject to shocks; this implies that the returns may exhibit fatter tails than the Gaussian. The literature of the suitability of different density functions as models of asset returns is dominated by applications to the equity market. We draw on these investigations to suggest alternatives to the Gaussian density to model uncertainty in oil returns. Osborne (1959) was one of several to have suggested that the behaviour of equity returns was consistent with Brownian motion and returns were Gaussian distributed. The Gaussian hypothesis is undermined by the evidence from many empirical analyses of asset returns. Mandlebrot (1963) argued that a stable distribution was more suitable than the Gaussian, based on observed leptokurtosis in asset return. Akgiray and Booth (1988) considered a stable-law model for individual US equity returns; they concluded that empirical tails were thinner than implied by the fitted stable distribution and preferred a skewed distribution with fatter tails than the Gaussian.

Here we consider density functions predominantly for conditional distributions of returns, thus ease of parameter estimation has to be considered. This means that the stable law will not be considered here (see McCulloch, 1986, for discussion of parameter estimation). The random variables considered here are discussed in the following list, for convenience the density functions and expressions for their excess kurtosis are given in the Appendix.

The Gaussian random variable is included as a benchmark. Since the densities are estimated given a mean and variance, no parameters are estimated for the Gaussian.

A mixture of Gaussian random variables captures leptokurtosis and can capture skewness if the means of the components are not identical. Using daily returns from thirty US stocks, Kon (1984) found that the log-likelihood of the mixture was greater than for the Student's t. Mixtures of two, three and four Gaussians were used. Melick and Thomas (1997) use option prices to recover the implied probability density of crude oil prices. They use a mixture of three Gaussians with different means and standard deviations. Venkataraman (1997) used a VaR (Value at Risk) approach to demonstrate that foreign exchange rate returns were more consistent with a mixture of two Gaussians than the Gaussian distribution. In general, *m* Gaussians can be mixed and the choice of *m* is decided by a likelihood ratio test. The number of estimated parameters is *3m-1* with two constraints to give the required mean and variance.

A mixture of Gaussian and Laplace random variables. The motivation behind the use of the Laplace is to use its thicker than Gaussian tail to model empirical leptokurtosis. There are five estimated parameters subject to two mean and variance constraints.

The Normal Inverse Gaussian random variable (NIG). The density function of this random variable has four parameters. Given the mean and variance of the data, this leaves two parameters free to describe the shape of the distribution. Barndorff-Nielson (1997) proposed it as a model of stock returns. Rydberg (1999) used this density function to model daily returns of US equities. Its use in risk analysis was demonstrated by Lillestøl (2000), who fitted the density function to returns on the S&P 500 and FT Actuaries indices. (Other transformations of the Gaussian include Tukey's *g* and *h* distributions, used by Mills (1995) to model daily returns on three London FTSE indices. Edgeworth-Sargan distributions are used by Mauleon and Perote (2000) to model stock market indices; the Gram-Charlier distribution is used by Verhoeven and McAleer (2004) for a NASDAQ index; these distributions use polynomials of Gaussian random variables.)

Student's t random variable.[2] Blattberg and Gonedes (1974) proposed the Student's *t* random variable as a model for the daily returns on US stocks. Comparing the fit of the stable and the Student's *t*, the Student's *t* was considered the better fit for two reasons. Firstly, the estimated degrees of freedom increased as the frequency of observation decreased, indicating a trend towards Normality, this is contrary to the assumption of the stable law where non-Normality would persist under addition.

---

[2] The Student's *t* and the GED are special cases of the Generalised *t* of McDonald and Newey (1988). Initially, we considered a further generalisation, the Skewed Generalised Student's *t* random variable, Theodossiou (1998) introduced this density function to model the empirical behaviour of financial time series. However, since this density was consistently less accurate than one or both of its nested special cases, it was dropped from the analysis.

Secondly, the log-likelihood of the Student's *t* was always greater than that of the stable.

Generalised error random variable (GED). This random variable has been used to capture leptokurtosis in the conditional returns in conjunction with ARCH-GARCH models by Nelson (1991) and in conjunction with stochastic volatility by Liesenfeld and Jung (2000).

Compound Poisson with Gaussian compounding. Tucker and Pond (1988) considered this random variable as a model for foreign exchange rates, its properties are derived in detail by Press (1968). The main reason for its inclusion in this study is Duan, Ritchken and Sun (2004, 2006) (*DRS*), who propose a GARCH model with jumps (modelled by a compound Poisson). The structure of this model is restricted to allow it to approximate a continuous process which can be used for derivative pricing via Monte Carlo simulation. Since either a departure from the Gaussian distribution or the introduction of stochastic volatility (via a GARCH model) or both tends to lead to market incompleteness, the ability of this model to be used for derivative pricing is attractive. Duan, Ritchken and Sun (2004) derive a general model that can be estimated using both asset price and option data, they also give a simplified version for asset price data alone, it is this model we will describe below. In contrast to the other density functions, this density function will not be used with the time series models in (12) to (15), we will only consider the *DRS* implementation of a restricted non-linear GARCH model with jumps. We adapt their notation where necessary to avoid confusion with terms used elsewhere in this paper.

The returns are defined as

$$Z_t = \sqrt{h_t}\,\bar{J}_t + \varpi_t \tag{16}$$

where h_t is a scale factor for the local, time-dependent, variance, $\bar{J}_t$ is a jump process and ϖ_t is a GARCH in mean term. The jump process is a sum of Gaussian random variables, defined as

$$\bar{J}_t = \bar{X}_t^{(0)} + \sum_{j=1}^{N_t} \bar{X}_t^{(j)} \tag{17}$$

where N_t follows a Poisson process with parameter λ; $\bar{X}_t^{(0)} \sim N(0,1)$ is the diffusion process $\bar{X}_t^{(j)} \sim N(\bar{\mu}, \bar{\gamma}^2)$ are the jumps. The local variance of returns is $h_t\left(1+\lambda\left(\bar{\mu}^2+\bar{\gamma}^2\right)\right)$ where the scale factor is assumed to follow a non-linear GARCH(1,1) process,

$$h_t = \alpha_0 + \alpha_1 h_{t-1}\left(\frac{\bar{J}_{t-1} - \lambda\bar{\mu}}{\sqrt{1+\lambda\left(\bar{\mu}^2+\bar{\gamma}^2\right)}} - c\right)^2 + \beta_1 h_{t-1} \qquad (18).$$

The GARCH in mean term is:

$$\varpi_t = \delta\sqrt{h_t} - \frac{1}{2} h_t\left(1+\lambda\left(\bar{\mu}^2+\bar{\gamma}^2\right)\right) \qquad (19).$$

The remaining details are given in the Appendix.

Although Nelson modelled the conditional distribution of asset returns, most of the above analyses (Kon, Venkataraman, Blattberg and Gonedes, Lillestøl, Theodossiou) refer to the unconditional distribution of asset returns. Modelling the conditional heteroscedasticity of asset returns captures some of the observed leptokurtosis in the unconditional distribution of returns. In his study of the dynamics of commodity prices and volatility, Pindyck (2004) suggests that short term volatility is, at least partially, driven by speculative noise trading, in addition to reactions to changed fundamentals. In the following analysis, the emphasis will be on the estimation of conditional density functions.

2.5 Model validation

Quasi-maximum likelihood is used throughout as the objective of the estimation algorithm. To both validate the estimation algorithm and to measure how well a density function captured the properties of data generated by another density function a simulation exercise was performed. GARCH time series were simulated using each of the seven density functions identified in Section 2.4 as error densities. Each of the resulting 210 (7 x 30) time series has 4000 observations (comparable with the observed oil price series). Each series was estimated using each of the seven density functions, the point of interest is how well the kurtosis of the generated series is estimated by the 'wrong' model. Our emphasis on kurtosis as the dominant departure from normality will be justified in Section 3. This experiment will provide evidence about the relative robustness of the estimated densities. If, for example, we found that a particular density function provided accurate estimates of the kurtosis of series, regardless of the density used for generation, then this would suggest that the estimating density was a good candidate for modelling the real data. The results are summarised in Table 1. The values of the excess kurtosis of the generated series are summarised in the first rows of the Table, the successive rows summarise the accuracy with which the excess kurtosis of the generated series is estimated by each density function, the mean error and the root mean square error of the estimates are given for each combination.

The mean error is $\sum_{j=1}^{J} \left(\gamma_{2,j} - \hat{\gamma}_{2,j}\right) \Big/ J$ and the root mean square error (*rmse*) is $\sqrt{\sum_{j=1}^{J} \left(\gamma_{2,j} - \hat{\gamma}_{2,j}\right)^2 \Big/ J}$, where γ_2 is the kurtosis of the generated series and $\hat{\gamma}_2$ is the estimate. Both terms are computed using the expressions for excess kurtosis given in the Appendix.

From the table, it can be seen that each generated density function is best estimated by itself as the estimating density function in terms of lowest *rmse*. The right hand column summarises the accuracy of each estimating density function; the mixture of two Gaussians has the lowest *rmse*, followed by DRS, and the Gaussian has the highest. In the summary rows at the bottom of the Table, it can be seen that the generated Gaussian was estimated with the least error. The mixture of two

Gaussians has the highest *rmse* implying that the other density functions do not easily match its kurtosis.

In summary, if the data generating process is a mixture of two Gaussians then the kurtosis is not well captured by the other densities considered. Perhaps more importantly, of the data generating processes considered, the mixture of two Gaussians estimates the kurtosis more accurately than the alternatives.

2.6 Comparison of density forecasts

The accuracy of the density forecasts following from the estimation of these densities will be the main criterion for judging which, if any, of those considered are appropriate for modelling oil prices. Here we discuss how this criterion can be applied. Measuring the accuracy of density forecasts of financial time series has been approached in several different ways. The approaches found in the literature are summarised in Table 2. The right hand column in the Table gives the accuracy measure used. The background to these measures is now described.

For a time series $\{Z_t\}$, we denote the true density of an observation at time t as $f_t(z_t \mid \Theta_{t-1})$ and a range of candidate density functions are denoted as $g_{it}(z_t \mid \Theta_{t-1})$ where i is an index identifying the candidate and Θ_{t-1} represents information available up to time *t-1*. The probability integral transformations of these densities are $F_t(z_t) = \int_{-\infty}^{z_t} f_t(u \mid \Theta_{t-1})\,du$ and $G_{it}(z_t) = \int_{-\infty}^{z_t} g_{it}(u \mid \Theta_{t-1})\,du$.

One approach to the comparison of density functions develops the Kolmogorov-Smirnov (KS) methodology for density comparison. The error measure used is

$$\text{Distance}(f, g_i) = E\left[\left(G_{it}(z_t) - F_t(z_t)\right)^2\right] \qquad (20)$$

Recent developments and the background to this approach are given by Corradi and Swanson (2006).

Another approach centres on the uniform distribution. Diebold, Gunther and Tay (1998) note that if candidate j is the true density, i.e. $g_{jt}(z_t \mid \Theta_{t-1}) \equiv f_t(z_t \mid \Theta_{t-1})$ then $G_{jt}(z_t)$ is a uniform U(0,1) random variable. Hong and Li (2005) and Hong, Li and Zhao (2007) develop this approach by devising a portmanteau statistic (analogous to the Box-Pierce-Ljung statistic) for $G_{it}(z_t)$ based on the density of the bivariate random variable $\left(G_{it}(z_t), G_{it-k}(z_{t-k})\right)$ (for $k \geq 1$) being a product of two uniform U(0,1) densities.

A third approach is based on the use of an information criterion. Berkowitz (2001) argues that testing a hypothesis of normality is more straightforward than testing for an arbitrary distribution. Thus he considers $Y_{it} = \Phi^{-1}\left(G_{it}(z_t)\right)$ where $\Phi(\ )$ is the standard normal cumulative density function. Some authors, see Rapach and Wohar (2006) use the Doornik and Hansen (1994) test for the normality of the Y_{it} (this test is based on a sum of transformed skewness and kurtosis statistics).

Here we follow the approach of Bao, Lee and Saltoglu (2004, 2007), they argue that the Kullbach Leibler information criterion (KLIC) should be used to measure the distance between actual and forecast densities:

$$\text{Distance}\left(f, g_i\right) = E\left[\left(\ln\left(f_t\left(z_t\right)\right) - \ln\left(g_{it}\left(z_t\right)\right)\right)\right] \qquad (21).$$

For a set of M observations of Z_t, we have

$$KLIC = \frac{1}{M}\sum_{t=1}^{M}\left(\ln\left(f_t\left(z_t\right)\right) - \ln\left(g_{it}\left(z_t\right)\right)\right) \qquad (22)$$

Since $f_t\left(z_t \mid \Theta_{t-1}\right)$ is unknown, Bao, Lee and Saltoglu follow Berkowitz's argument and use the KLIC to measure the departure of Y_t from being independently N(0,1). The KLIC in this case is

$$KLIC^* = \frac{1}{M}\sum_{t=1}^{M}\left(\ln\left(\frac{1}{\hat{\sigma}}\phi\left(\frac{y_t - \hat{\mu} - \hat{\rho} y_{t-1}}{\hat{\sigma}}\right)\right) - \ln\left(\phi\left(y_t\right)\right)\right) \qquad (23)$$

where $\phi(\ )$ is the standard normal density and $\hat{\mu}, \hat{\sigma}, \hat{\rho}$ are respectively the sample mean, standard deviation and first order correlation of Y_t. Mitchell and Hall (2005) advocate the use of KLIC as a "unified statistical tool" for the evaluation and comparison of density forecasts. Berkowitz's (2001) objective was to construct a test for the hypothesis that $g_{jt}\left(z_t \mid \Theta_{t-1}\right) \equiv f_t\left(z_t \mid \Theta_{t-1}\right)$, under this hypothesis, using (23), 2.M.KLIC* is a χ_3^2 random variable.

3 The data and an initial analysis

For our detailed analysis of the market, we use prices from April 1991 onwards. During this period the market is liquid and the price spike caused by the Iraqi invasion of Kuwait is avoided. The data set consists of daily prices from 8 April 1991 to 10 December 2008 for Brent (Current Month FOB US$/bbl) and WTI (Spot Cushing US$/bbl) obtained from Datastream.

Price and a twenty working day volatility series are plotted in Figure 2 for both data sets. As one might expect, the oil price series are similar, although Brent appears to have more distinct peaks in volatility than WTI. Normally it would be appropriate to present stylised facts describing the time series of returns such as skewness and excess kurtosis, however as mentioned earlier, these price series are subject to shocks, possibly the output of a jump-diffusion process. In these circumstances statistics describing higher moments are likely to be unduly influenced by outlying observations. In order to discover the magnitude of the diffusion and the relative frequency and magnitude of jumps, we compute a 260 day rolling inter-quartile range (IQR) and look at each day's return, z_t, in the context of the most recent IQR. We define the magnitude of an outlier by this function $Out(Z_t)$:

$$Out\left(Z_t\right)=\begin{cases}\left(\dfrac{Z_t-UQ_t}{IQR_t}\right) & if\ Z_t>UQ_t+2IQR_t \\ \left(\dfrac{Z_t-LQ_t}{IQR_t}\right) & if\ Z_t<LQ_t-2IQR_t \\ 0 & otherwise\end{cases} \qquad (24)$$

where UQ_t , LQ_t and IQR_t are the upper quartile, lower quartile and inter-quartile range of the returns $\left(Z_k,\quad k=t,t-1,\ldots,t-260\right)$. Rolling upper and lower quartiles of returns and the outliers for both data sets are shown in Figure 3. Discussing diffusion first and taking the inter-quartile range as a guide to longer term volatility, we see that this volatility increased for both oil series until 1999 after which the inter-quartile range seems stable. The post 1999 inter-quartile range of daily returns for both oils is around 0.03, this is equivalent to an annualised standard deviation of 36% p.a. (making the Gaussian assumption that the inter-quartile range is 1.35 standard deviations). Secondly, we discuss jumps as represented by the outliers. Both oil series exhibit clusters of outliers in 1993, 1994, 1996 and 1998 (with at least 14 outliers per year), after that the frequency of jumps diminishes until 2008 where there are 19 outliers for Brent, 23 for WTI, between September 1st and December 10th. To be visible on the graph, the outlier needs to be a distance of two inter-quartile ranges from the relevant quartile, a normal random variable would achieve this once in five years, for three inter-quartile ranges the frequency would be once in 6000 years, for four IQRs the frequency would be once in two million years. The severity of the jumps is lower for Brent with 8 outliers greater than four IQRs and 2 outliers greater than six IQRs; the corresponding frequencies for WTI are 17 and 6.

Both data sets are subject to outliers that are sufficiently large that they are unlikely to be captured by Gaussian models even if the heteroscedasticity visible in Figure 2 is included in the modelling process.

4. Analysis of Daily Brent and WTI Crudes 1991 to 2008

In this section, we examine the ability of the density functions identified in Section 2.4 to capture the behaviour of the returns of both oil series used with the time series model described in equations (12) to (15). This analysis is comprised of model estimation, generation of forecasts, evaluation of density forecasting accuracy and discussion of the evidence for mean reversion. The structure of the time series of oil price returns is examined first. Some properties, the presence of auto-regressive and GARCH terms are established at the estimation stage. The property of mean reversion is left open up to the forecasting stage by considering models with and without this property.

The experimental design to measure the accuracy of density forecasts for horizons of up to a year involves several compromises. The size of the data set permits only twelve non-overlapping one year ahead forecasts; this is too few to evaluate the accuracy of the forecast density. If overlapping forecasts are used then the forecasts become auto-correlated which means that only the coverage of the density forecasts can be measured and tested. Another practical consideration is the computation time of repeated model estimation and simulation for the forecast density. It was decided to generate overlapping one year ahead forecasts in order to measure density coverage at a cost of not measuring the correlation between successive forecasts.

We denote the parameter estimates for a candidate model using data available from $t = 0,1,\ldots,T$ as a vector, $\boldsymbol{\theta}_T$. Each candidate model was fitted to the data over an increasing length of time, starting at five years (T = 1300). The estimation region is repeatedly increased by 10 working days steps up to a maximum of 17 years 7 months (T = 4600).

We will discuss the basic time series properties first; then the generation of forecasts; then the evidence for the choice of density function and the evidence for mean reversion.

4.1 Model estimation

Although Gaussian models with constant variance (discrete Brownian motion) are included in the forthcoming evaluation of density forecasts, the significance of GARCH effects in both return series for all models considered is pervasive.

Given the number of outliers in the return data for both series and the heteroscedasticity captured by the GARCH process, the auto-correlation function of returns would be an unreliable indication of any autoregressive process in returns. Thus it was decided to use the autocorrelation of standardised one step ahead forecast errors as the criterion for including an AR process in the model of returns. The standardised returns are $\left(\frac{\hat{\varepsilon}_t}{\sqrt{\hat{\sigma}_t^2}} \right)$ where $\hat{\varepsilon}_t = Z_t - E\left(Z_t \mid \Theta_{t-1}\right)$ and the model coefficients (using a Gaussian density function) were re-estimated every ten observations (Θ_t represents the information available at time *t*). An analysis of the basic time series models for both return series is given in Table 3. Asymmetry in volatility was detected in the WTI but not detected in the Brent return series. In studies of equity returns, price falls tend to generate higher volatility than price rises ($\xi > 0$), for WTI $\xi < 0$ indicating that a price rise tends to lead to increased volatility.

The values of the excess kurtosis of the standardised one step ahead errors are positive for both series, indicating that GARCH alone does not capture the excess kurtosis in the return series and that a fatter tailed random variable than the Gaussian is needed. For both oil series, using the Ljung Box test, there is no evidence of auto-correlation in the error series, implying no need for any autoregressive terms in the model.

The Gaussian density was replaced by each of the six non-Gaussian densities in the estimation of the model for both return series defined in Table 3. The first five of these models were estimated without and with mean reversion and without asymmetric GARCH where appropriate. The DRS model estimated is described by

(17), (18) and (19). The motivation for the estimation of these extra models is to ascertain whether these features improve density forecasting accuracy.

4.2 Forecast generation, measurement and comparison of density forecasting accuracy

Here we describe how the density forecast is computed and then explain the computation of the Berkowitz statistic. From the data for $t = 0,1,\ldots,T$, estimation yields the expected values for the model parameters, a vector $E\left(\boldsymbol{\theta}_T \mid \Theta_T\right)$ and a covariance matrix $V\left(\boldsymbol{\theta}_T \mid \Theta_T\right)$ which describes the estimation error. At each origin, *T*, simulation is used to generate the density of returns over all horizons, *h*, (*h* = 1, …., *H*) up to one year ahead (H = 260 days). Note that there are two sources of randomness in this simulation: estimation error and the noise process. At each origin, *T*, *K* simulation replications are generated (we used K = 1000 simulations); for simulation *k* (*k*=1,.., *K*), a sample value of parameters, $\tilde{\boldsymbol{\theta}}_{k,T}$, is drawn from $N\left(E\left(\boldsymbol{\theta}_T \mid \Theta_T\right); V\left(\boldsymbol{\theta}_T \mid \Theta_T\right)\right)$. The simulated return is $\tilde{Z}_{k,T+h} = g\left(\tilde{\boldsymbol{\theta}}_{k,T}, \varepsilon_{k,T+j}\; j = 1,\ldots,h\right)$ where *g*() is defined by the time series equations (12) to (15) and the density function. The simulated price is $\tilde{S}_{k,T+h} = S_T \exp\left(\sum_{j=1}^{h} \tilde{Z}_{k,T+j}\right)$. The set of simulated prices for a given origin and horizon is sorted to ascending order $\left\{\tilde{S}'_{1,T+h},\ldots,\tilde{S}'_{K,T+h}\right\}$ and the predicted cumulative density of the price S_{T+h} is $\hat{F}\left(S_{T+h} \mid \Theta_T\right) = \frac{k'}{K}$ where $S_{T+h} \geq \tilde{S}'_{k',T+h}$ and $S_{T+h} < \tilde{S}'_{k'+1,T+h}$.

The Berkowitz statistic (proportional to KLIC) is used as the primary measure of accuracy. Due to the nature of the experiment, where forecasts are prepared for up to 260 days every 10 days, forecasts for more than 10 days ahead will be correlated, thus we will primarily test the coverage hypothesis. The Berkowitz statistic is calculated cumulatively as the latest forecast origin moves towards the end of the data set and for different horizons.

Modifying the expression in (19), the statistic is given below.

$$\text{Berkowitz statistic} = 2\sum_{t=1}^{M}\left(\ln\left(\frac{1}{\hat{\sigma}}\phi\left(\frac{y_{t+h}-\hat{\mu}}{\hat{\sigma}}\right)\right) - \ln\left(\phi\left(y_{t+h}\right)\right)\right) \qquad (25)$$

where $y_{t+h} = \Phi^{-1}\left(\hat{F}\left(S_{t+h} \mid \Theta_t\right)\right)$. For each model, the Berkowitz statistic is calculated for a range of origins (*M*) and for a range of horizons (*h*). The values of the horizon, *h*, run from 10 (2 weeks) through to 260 (a year).

To give an initial impression of the accuracy of the density forecasts of the different models over different horizons, a summary of the Berkowitz statistics with the last forecasting origin at 10 December 2007 is given in Table 4 for both oil series. We show results for each density function without and with constant mean reversion, except DRS where one version only is considered. For two density functions (mixture of two Gaussians and Student's t) we also show the results for reversion to an adaptive mean. Viewing the statistics as a scaled KLIC, lower values indicate greater accuracy, i.e. a smaller distance between the observed and the fitted density function. Viewing the statistics as a likelihood ratio test, under the null hypothesis that the density function fitted is the 'true' density, the critical values at 5% and 1% are $\chi^2_{2,0.95} = 6.0$, $\chi^2_{2,0.99} = 9.2$ respectively.

The most obvious theme in these results is that in general the accuracy of the density forecasts deteriorates as the horizon lengthens; for example nearly all the models are 'acceptable' for a ten day horizon for Brent and WTI (with statistics less than 9.2) but nearly a third of these densities can be rejected at a 30 day horizon, nearly two thirds by a 70 day horizons and 90% for 130 days and beyond. We consider first the arbitrage pricing models described in Section 2.1, equations (1) and (2). The hypothesis of Brownian motion (represented by the Gaussian density with constant variance) can be rejected for Brent, but cannot be rejected for short horizons (up to 30 days ahead) for WTI. Indeed, the density forecasts for a year horizon from Brownian motion are more accurate than many of the alternatives. The hypothesis of mean reversion with constant variance is rejected for both series, it is only supportable for the shortest horizon for WTI. Although the asymmetry in the variance process for WTI became insignificant in the presence of non-Gaussian density functions; its inclusion for WTI with the Gaussian density (with no mean reversion) does lead to an improvement in density forecast accuracy for all horizons. The inclusion of constant mean reversion in the model leads to a deterioration in accuracy in nearly all cases for Brent and WTI.

Let us investigate the dominance of density forecasts of models with no mean reversion over the corresponding model with constant mean reversion. In Table 5 we show the evolution of the estimated parameters of a mean reversion model with a Gaussian density and GARCH. The long term mean is converted from the log form used in (12) and the reversion rate is given. The mean reversion rate is significantly greater than zero until April 2004 for Brent. The log price and its long term mean are shown in Figure 4 (we shall discuss the adaptive mean also shown later). We see that the log price crossed its long term mean in December 2001 and has not returned to it since. The estimated value of the reversion rate is at its highest using data up to 1998, the rate decreases thereafter.

Returning to the measures of density forecasting accuracy in Table 4, let us consider the models without mean reversion in more detail. Looking at the non-Gaussian density functions considered for their ability to capture the fat tails observed in both series, we see the following. For both series, three density functions stand out, in decreasing order of performance these are: the Gaussian mixture, Student's *t* and DRS, they exhibit the lower one year horizon Berkowitz statistics. For Brent, these statistics are acceptable at all the horizons considered; for WTI, the statistics become significant at some intermediate horizons. For the three other densities; the Gaussian-Laplace mixture, the Normal Inverse Gaussian and the generalised error distribution; for WTI, the Berkowitz statistics are acceptable up to a 30 day horizon; for Brent, they are acceptable up to a 70 day horizon. Henceforth, since we are concerned with density forecasting accuracy for horizons of a year or more, we will focus on the Gaussian mixture, Student's t and the DRS. The parameter estimates for these densities are given in Table 6 for both series.

In Table 7, we show the summary statistics of the one to ten day ahead standardised forecast errors from these two models. The means are effectively zero and the standard deviations are slightly above one, showing a slight tendency to underestimate volatility. Following the discussion in Section 2.5 about the ability of the density functions to capture kurtosis, a comparison between the observed excess kurtosis of the forecast errors and the median of the excess kurtosis implied by the density estimates is carried out. The number of estimated parameters required to define excess kurtosis by each density function is given in parentheses. The comparison shows that DRS (3 parameters) provides the closest estimates, followed by the mixture of two Gaussians (2 parameters) which provides a far closer estimate of kurtosis than the Student's t (1 parameter).

A possible conjecture from the earlier discussion of the constant mean reversion models is that reversion may be occurring but to a shorter term mean than one based on the whole estimation region. To investigate this conjecture, we estimate the model of reversion to a time dependent mean given in (14) and (15) using the mixture of two Gaussians and the Student's t densities. In (15), the parameter, γ_t , is time varying in the sense that different values are appropriate at different time periods. The mean of ln(price) changes little over the period 1991 to 1999 thus γ_t would be very small; from 1999 onwards the log(price) starts trending upwards, requiring a larger value for γ_t. To allow for this within the rolling estimation procedure, where the origin is advanced in 10 day steps, the parameter γ_t is estimated over the last three years (780 days) of data, the long term mean prior to this comes from previous estimations. That is, for the estimation region t=1,,,,K, γ_t is optimised using the observations for K-779 to K, the values for η_t for t=1,.., K-780 are taken from the previous estimation using the region t=1,..,K-10.

The behaviour of the adaptive mean (using a mixture of two Gaussians) for Brent is shown with a running estimate of the long term mean in Figure 4. The log price crosses the adaptive mean several times after December 2001 (the last time it crossed the long term mean). In this case the reversion rate remains significant until July 2005 when it falls below 0.004. Even with an adaptive mean, the evidence that prices revert to this mean only persists fifteen months longer than the evidence of reversion to a constant mean.

Thus for horizons of up to one year, we find that the most accurate density forecasts are produced by a model without mean reversion with either the mixture of two Gaussians, the Student's t or DRS. In addition to their accuracy, the hypotheses that the data generating process is consistent with the models used are acceptable (for all horizons for Brent prices, for most horizons for WTI). The accuracy of the DRS is important as its structure can be used via Monte Carlo methods to value contingent claims on oil prices. Although outside of the scope of this paper, with the use of option data, the DRS model can be generalised further to include the valuation of options across a range of strike prices (see Duan, Ritchken and Sun, 2004).

4.3 Density forecasting accuracy for beyond one year

Clearly the inferences we can draw from our analyses depend on the data used for estimation. Whereas the data available to Smith and McCardle (1999) led them to accept the hypothesis of mean reversion, our extra decade of data allows us to reject the hypothesis of reversion to a long term mean. A possible conclusion for future modelling is that one should monitor a portfolio of models including those with reversion to an adaptive mean.

Having established that the combinations of a symmetric GARCH model with no mean reversion and either the mixture of two Gaussians, the Student's t densities or DRS produces more accurate density forecasts for both oil price series for up to one year ahead, the obvious question is for how long a horizon does this accuracy persist. To answer this question, these models were used to forecast up for longer horizons and the accuracy of the density forecasts measured by the Berkowitz Statistic. In Figure 5, the Berkowitz Statistic for Brent and WTI prices modelled by the mixture of two Gaussians is shown for horizons up to two and a half years ahead. The statistic is shown using the most recently available data for each horizon for estimation (24 November 2008 for 10 days ahead, 10 November for 20 days ahead, etc). For comparison, we show corresponding results available one and two years earlier. The accuracy for longer horizons has deteriorated over the last two years. For Brent (WTI), the mixture of two Gaussians was a plausible model up to a horizon of 490 (460) days (using data up to 27 November 2006); comparable horizons fall to 450 (430) days (26 November 2007) and 410 (370) days using the most recent data.

In order to determine the cause of the deterioration in density forecast accuracy, histograms are shown of the empirical cumulative density functions (cdf) for different horizons (for Brent, mixture of two Gaussians) in Figure 6. If the density is correctly forecast the histogram should be uniform (with a frequency of 5%); if the mean tends to be under- (over-)estimated, a mode should appear for cdf values greater (lesser) than 0.5; if the variance is under- (over-)estimated the density will be U shaped (peaked in the middle). From Figure 6, we see that the cdfs are approximately uniform for a year ahead. For two years ahead, where we have seen in Figure 5 that the accuracy of the density forecasts begins to deteriorate, there is evidence that the mean price has been underestimated. For three years ahead, there is evidence of both underestimation of the mean price and overestimation of variance.

For the evaluation of oil investments over longer horizons of a decade or more, these findings indicate that the accuracy of estimated probabilities will deteriorate two or more years into the future. Depending on the nature of the project, this may mean that a decision to increase or suspend investment in the near term may be relatively well informed. The decreasing accuracy of probability estimates in the longer term will lead to less accurate estimates of the project's lifetime value.

4.4 Examples of one year ahead forecasts

We demonstrate the density forecasts using two origins, 26th November 2007 and 24th November 2008. In the first case, the forecasts can be compared with actual data, in the second case the data are not available at the time of writing. In Figure 7, for the first origin we show the predicted development of the price density for Brent using the mixture of two Gaussians via a selection of percentiles. The year in question includes a peak of US$140 in July; however by the end of the year modelled the price has fallen to US$50. As remarked in Section 3, the return series for this year exhibits an unusually high number of outliers. On the right of Figure 7, there is the predicted density function; the cdf for the actual price is 2.5%. In Figure 8, the predictions for the currently unknown future year ending 23rd November 2009 is given. The conditional probability of the price remaining under US$100 is 80%.

5 Summary and Conclusions

We are concerned with modelling oil price risk over a horizon of a year or longer, to provide information about risk for the evaluation of real options where oil price is a risk factor. The accuracy of density forecasting by candidate models is used as our main criterion as point forecasting alone gives no information about risk.

Our evaluation of the two modelling approaches suggested by the arbitrage pricing literature, geometric Brownian motion and mean reversion, showed that both models ceased to be plausible for horizons of three months or less. Geometric Brownian motion is not supportable as a long term model because of time varying volatility and the returns exhibit many jumps of a magnitude completely inconsistent with a Gaussian density function. The time varying volatility is captured by a GARCH process. We investigated several alternative non-Gaussian densities and found that a mixture of two Gaussians was most successful at capturing the jump diffusion process, providing plausible density forecasts up to a year ahead. The Student's t density is also a plausible model but did not estimate the observed kurtosis of the returns as well as the Gaussian mixture; this finding was consistent with our model validation exercise which demonstrated the flexibility of the Gaussian mixture in the estimation of kurtosis for a variety of data generating processes.

Adding mean reversion to the model does not help. Reversion to a constant mean is shown to have ceased being a plausible model in 2004. We investigated reversion to an adaptive mean but this hypothesis was only plausible for a further eighteen months. In general, inclusion of mean reversion led to a deterioration in the accuracy of density forecasting.

Our investigation shows that oil price behaviour can be modelled as a jump-diffusion process with time varying volatility and no mean reversion. We regard the structural breaks that occur in the oil market too frequent to model and consider them as jumps, indeed evidence from the literature suggests that modelling breaks does not improve out of sample forecasting accuracy. The accuracy of the density forecasts for the Gaussian mixture for both Brent and WTI was found to be plausible for horizons just less than two years. Beyond two years, we found that the mean price tends to be underestimated and the volatility over estimated. We recognise that many real option appraisals will consider projects where the horizons are greater than two years. We summarise the implications thus: accurate probability estimates in the short term will facilitate decisions to continue or suspend investment; in the longer term poorer probability estimates will make the lifetime value of a project less accurate.

References

Akgiray, V. and Booth, G.G., 1988, The Stable Law Model of Stock returns, *Journal of Business & Economic Statistics*, 6, 51-57.

Bao, Y., Lee, T-H. and Saltoglu, B., 2004, A test for density forecast comparison with applications to risk management, Working Paper, Dept. of Economics, University of California, Riverside.

Bao, Y., Lee, T-H. and Saltoglu, B., 2007, Comparing density forecast models, *Journal of Forecasting*, 26, 203 – 225.

Barndorff-Nielsen, O.E., 1997, Normal Inverse Gaussian process and stochastic volatility modelling, *Scandinavian Journal of Statistics*, 24, 1-13.

Bauwens L, Giot P, Grammig J, Veredas D, 2004, A comparison of financial duration models via density forecasts, *International Journal of Forecasting*, 20, 589-609.

Berkowitz, J., 2001, Testing density functions with applications to risk management, *Journal of Business & Economic Statistics*, 19, 465 – 474.

Blattberg, R.C. and Gonedes, N.J., 1974, A comparison of the stable and Student distributions as statistical models for stock prices, *Journal of Business*, 47,244-80.

Brennan, M.J. and Schwartz, E.S., 1985, Evaluating natural resource investments, *The Journal of Business*, 58, 135 – 157.

Cabedo, J.D. and Moya, I., 2003, Estimating oil price 'Value at Risk' using the historical simulation approach, *Energy Economics*, 25, 239 – 253.

Clements, M.P. and Hendry, D.F., 1996, Intercept corrections and structural change, *Journal of Applied Econometrics*, 11, 475-494.

Corradi V, Swanson N.R., 2006, Predictive density and conditional confidence interval accuracy tests, Journal of Econometrics 135, 187-228.

Cortazar, G. and Naranjo, L., 2006, An N-factor Gaussian model of oil futures prices, *The Journal of Futures Markets*, 26, 243 – 268.

Cortazar, G. and Schwartz, E.S., 1997, Implementing a real option model for valuing an undeveloped oil field, *International Transactions in Operations Research*, 4, 125 – 137.

Cortazar, G. and Schwartz, E.S., 1998, Monte Carlo Evaluation Model of an Undeveloped Oil Field, *Journal of Energy Finance & Development*, 3, 73-84.

Deaton, A. and Laroque, G., 1992, On the behavior of commodity prices, *The Review of Economic Studies*, 59, 1 – 23.

De Gooijer, J.G., and Zerom, D., 2000, Kernel based multistep-ahead predictions of the US short-term interest rate, *Journal of Forecasting*, 19, 335 – 353.

Diebold F.X., T.A. Gunther and A.S. Tay, 1998, Evaluating density forecasts with applications to financial risk management, *International Economic Review*, 39, 863 – 883.

Doornik, JA, and Hansen, H., 1994, A practical test for univariate and multivariate normality, Discussion paper, Nuffield College, Oxford.

Duan, J-C., Ritchken, P. and Sun, Z., 2006, Approximating GARCH-Jump Models, Jump-Diffusion Processes, and Option Pricing, *Mathematical Finance*, 16, 21–52.

Duan, J-C., Ritchken, P. and Sun, Z., 2004, Jump-starting GARCH: Pricing and Hedging Options with Jumps in Returns and Volatilities, Working Paper, University of Toronto and Case Western Reserve University.

Egorov A.V., Hong Y.M., Li H.T., 2006, Validating forecasts of the joint probability density of bond yields: Can affine models beat random walk? *Journal of Econometrics*, 135, 255-284.
Higham, D.J., 2001, An algorithmic introduction to numerical simulation of stochastic differential equations, *SIAM Review*, 43, 525 – 546.
Hong, Y. and Li, H., 2005, Nonparametric Specification Testing for Continuous-Time Models with Applications to Term Structure of Interest Rates, *The Review of Financial Studies*, 18, 37 - 84.
Hong Y.M., Li H.T., Zhao F., 2007, Can the random walk model be beaten in out-of-sample density forecasts? Evidence from intraday foreign exchange rates, *Journal of Econometrics*, 141, 736 - 776.
Hung, J-C., Lee, M-C. and Liu, H-C., 2007, Estimation of value-at-risk for energy commodities via fat-tailed GARCH models. *Energy Economics*, 30, 1173 -1191.
Kon, S.J., 1984, Models of stock returns - a comparison, *Journal of Finance*, 39,147-165.
Kumar, M.S., 1992, The Forecasting Accuracy Of Crude-Oil Futures Prices, *International Monetary Fund Staff Papers*, 39, 432-461.
Labys, W.C., and Granger, C.W.J., 1971, Speculation, Hedging and Commodity Price Forecasts, Heath-Lexington, Mass.
Liesenfeld R. and Jung, R.C., 2000, Stochastic volatility models: conditional normality versus heavy-tailed distributions, *Journal of Applied Econometrics*, 15, 137-160.
Lillestøl, J., 2000, Risk analysis and the NIG distribution, *Journal of Risk*, 2/4, 41-56.
Mandlebrot, B., 1963, The variation of certain speculative prices, *Journal of Business*, 36, 394-419.
Mauleon, I. and Perote, J., 2000, Testing densities with financial data: an empirical comparison of the Edgeworth-Sargan density to the Student t, *The European Journal of Finance*, 6, 225-239.
McCulloch, J.H., 1986, Simple consistent estimators of stable distribution parameters, *Communications in Statistics - Simulation and Computation*, 15, 1109 - 1136
McDonald, J.B. and Newey, W.K., 1988, Partially adaptive estimation of regression models via the generalised t distribution, *Econometric Theory*, 4, 428-457.
Melick, W.R. and Thomas, C.P., 1997, Recovering an asset's implied PDF from option prices: an application to crude oil during the Gulf crisis, Journal of Financial and Quantitative Analysis, 32, 91 – 115.
Mills, T.C., 1995, Modelling skewness and kurtosis in the London Stock Exchange – FTSE index return distributions, *The Statistician*, 44, 323-332.
Mitchell, J. and Hall, S.G., 2005, Evaluating, comparing and combining density forecasts using the KLIC with an application to the Bank of England and NIESR 'fan' charts of inflation, *Oxford Bulletin of Economics and Statistics*, 67, Supplement, 995 – 1033.
Nelson, D.B., 1991, Conditional Heteroskedasticity in Asset Returns, *Econometrica*, 59, 347-370.

Osborne, M.F.M., 1959, Brownian motion in the stock market, *Operations Research*, 7, 145-173. (Reprinted 1964 in The Random Character of Stock Market Prices, ed P. Cootner, MIT Press, Cambridge, Mass., 100-128.)

Panas, E. and Nini, V., 2000, Are oil markets chaotic? A non-linear dynamic analysis, *Energy Economics*, 22, 549 – 568.

Pindyck, R.S., 2004, Volatility and commodity price dynamics, *The Journal of Futures Markets*, 24, 1029 – 1047.

Press, S.J., 1968, A Modified Compound Poisson Process with Normal Compounding, *Journal of the American Statistical Association*, 63, 607- 613.

Rapach D.E., Wohar M.E., 2006, The out-of-sample forecasting performance of nonlinear models of real exchange rate behaviour, I*nternational Journal of Forecasting*, 22, 341-361.

Rydberg, T.H., 1999, Generalised hyperbolic diffusion processes with applications in finance, *Mathematical Finance*, 9, 183-201.

Schittenkopf, C., Dorffner, G., and Dockner, E.J., 2000, Forecasting time-dependent conditional densities: a semi-non-parametric neural network approach, *Journal of Forecasting*, 19, 355 – 374.

Schwartz, E.S., 1997, The stochastic behaviour of commodity prices: implications for valuation and hedging, *The Journal of Finance*, 52, 923 – 973.

Schwartz, E.S. and Smith, J.E., 2000, Short-term variation and long term dynamics in commodity prices, *Management Science*, 46, 893 – 911.

Smith, J.E. and McCardle, K.F. 1999, Options in the real world: lessons learned in evaluating oil and gas investments. *Operations Research*, 47, 1-15.

Stevens, P., 2005, Oil markets, *Oxford Review of Economic Policy*, 21, 19 – 42.

Theodossiou, P., 1998, Financial data and the skewed generalised t distribution, *Management Science*, 44, 1650-1661.

Tucker, A.L and Pond, L., 1988, The Probability Distribution of Foreign Exchange Price Changes: Tests of Candidate Processes, *Review of Economics and Statistics*, 70, 638-647.

Venkataraman, S., 1997, Value at risk for a mixture of Normal distributions: The use of quasi-Bayesian estimation techniques, *Federal Reserve Bank of Chicago Economic Perspectives*, (March/April), 2-13.

Verhoeven, P., and McAleer, M., 2004, Fat tails and asymmetry in financial volatility markets, *Mathematics and Computers in Simulation*, 64, 351 – 361.

Weigend, A. S., and Shi, S., 2000, Predicting daily probability distributions in S&P 500 returns, *Journal of Forecasting*, 19, 375 – 392.

Ye, M., Zyren, J., and Shore, J., 2006, Forecasting short-run crude oil prices using high- and low-inventory variables, *Energy Policy*, 34, 2736 – 2743.

Appendix. Density functions used in the analysis

The return random variable X has $E(X) = \mu$ and $V(X) = \sigma^2$. The excess kurtosis expressions are given for $\mu = 0, \sigma^2 = 1$

Random Variable; Excess Kurtosis; Density function; Comments
Gaussian: Excess kurtosis = 0; $f(x) = \frac{1}{\sigma}\phi\left(\frac{x-\mu}{\sigma}\right)$; $\phi(t) = \frac{e^{-t^2/2}}{\sqrt{2\pi}}$
Mixture of Gaussians: Excess kurtosis = $3\left(\lambda_1\sigma_1^4 + \frac{\left(1-\lambda_1\sigma_1^2\right)^2}{1-\lambda_1} - 1\right)$
$f(x) = \sum_{j=1}^{m}\frac{\lambda_j}{\sigma_j}\phi\left(\frac{x-\mu_j}{\sigma_j}\right)$; $\sum_{j=1}^{m}\lambda_j = 1$; $\lambda_j > 0$; $\sum_{j=1}^{m}\lambda_j\mu_j = \mu$ and $\sum_{j=1}^{m}\lambda_j\sigma_j^2 = \sigma^2$. We use m = 2.
Mixture of Gaussian and Laplace:
Excess kurtosis = $3\left(\lambda_1\sigma_1^4 - 2\frac{\left(1-\lambda_1\sigma_1^2\right)^2}{\left(1-\lambda_1\right)} - 1\right)$;
$f(x) = \frac{\lambda_1}{\sigma_1}\phi\left(\frac{x-\mu_1}{\sigma_1}\right) + \frac{\lambda_2}{2\sigma_2}e^{-\left\|\frac{x-\mu_2}{\sigma_2}\right\|}$; $\sum_{j=1}^{2}\lambda_j = 1$; $\lambda_j > 0$;
$\sum_{j=1}^{2}\lambda_j\mu_j = \mu$ and $\lambda_1\sigma_1^2 + 2\lambda_2\sigma_2^2 = \sigma^2$
Normal Inverse Gaussian: Excess kurtosis = $3\frac{\left(1+4\beta^2/\alpha^2\right)}{\delta\sqrt{\alpha^2-\beta^2}}$
$f(x) = A(\alpha,\beta,\eta,\delta)K_1\left(\delta\alpha q\left(\frac{x-\eta}{\delta}\right)\right)\exp(\beta x)q^{-1}\left(\frac{x-\eta}{\delta}\right)$
$q(z) = \sqrt{1+z^2}$; $A(\alpha,\beta,\eta,\delta) = \frac{\alpha\exp(\delta\gamma - \beta\eta)}{\pi}$;
$\gamma = \sqrt{\alpha^2-\beta^2}$ and $K_1(\ )$ is a modified Bessel function of the third kind of order 1. Given the mean and the variance of the returns, specifying α and β determines the other two parameters.
The variance of X is given by $V(X) = \sigma^2 = \delta\frac{\alpha^2}{\gamma^3}$ which determines δ.

Random Variable; Excess Kurtosis; Density function; Comments

Normal Inverse Gaussian (continued):

The expected value of X is $E\left(X\right)=\mu=\eta+\delta\dfrac{\beta}{\gamma}$ which determines η.

Student's t: Excess kurtosis = $\dfrac{6}{\left(\nu-4\right)}$;

$$f\left(x\right)=\frac{\Gamma\left(\frac{\nu+1}{2}\right)}{B\Gamma\left(\frac{\nu}{2}\right)\sqrt{\pi\nu}}\left(1+\frac{\left(\frac{x}{B}\right)^{2}}{\nu}\right)^{\frac{-(\nu+1)}{2}}$$

where ν is the degrees of freedom and the variance is $\left(\dfrac{\nu}{\nu-2}\right)B^{2}$.

Generalised Error Distribution

Excess kurtosis = $\dfrac{\Gamma\left(1/k\right)\Gamma\left(5/k\right)}{\Gamma\left(3/k\right)^{2}}-3$;

$$f\left(x\right)=\frac{k\exp\left(-\frac{1}{2}\left|\frac{x}{C}\right|^{k}\right)}{C\Gamma\left(1/k\right)2^{\left(1+1/k\right)}}$$

where $C=\sqrt{2^{-2/k}\dfrac{\Gamma\left(1/k\right)}{\Gamma\left(3/k\right)}}$ and k > 0

Compound Poisson with Gaussian compounding

Excess kurtosis = $\dfrac{\lambda\left(\bar{\mu}^{4}+6\bar{\mu}^{2}\bar{\gamma}^{2}+3\bar{\gamma}^{4}\right)}{\left(1+\lambda\left(\bar{\mu}^{2}+\bar{\gamma}^{2}\right)\right)^{2}}$

Denote $y_t=\sqrt{h_t}\,\bar{J}_t$, the density function of y_t is:

$$f\left(y_t\mid h_t,y_{t-1}\right)=\sum_{i=0}^{\infty}e^{-\lambda}\frac{\lambda^{i}}{i!}\phi\left(\mu_i\left(t\right),\sigma_i^{2}\left(t\right),y_t\right)$$

where $\phi\left(\mu,\sigma^{2},y\right)$ is the Gaussian density function, $\mu_i\left(t\right)=i\bar{\mu}\sqrt{h_t}$ and $\sigma_i^{2}\left(t\right)=h_t\left(1+i\bar{\gamma}^{2}\right)$.

Table 1. Summary of the excess kurtosis of series generated by seven density functions and the accuracy of the estimates of kurtosis by different density functions

Model used for time series estimation		Model used for time series generation: Gaussian	Mixture of 2 Gaussians	Gaussian Laplace mixture	Normal Inverse Gaussian	Scaled Student-t	General error distn.	DRS	Summary
	(1)	0	2.13	2.06	1.96	1.98	2.00	1.97	1.73
	(2)	0	0.70	0.45	0.12	0.09	0.15	0.35	0.79
Gaussian	(3)	0	2.13	2.06	1.96	1.98	2.00	1.97	1.73
	(4)	0	2.24	2.11	1.96	1.99	2.01	2.00	1.90
Mixture of 2 Gaussians	(3)	-0.03	-0.08	0.30	0.22	0.42	0.39	0.34	0.22
	(4)	0.06	0.38	0.43	0.38	0.60	0.55	0.41	0.43
Gaussian Laplace mixture	(3)	1.06	2.29	-0.10	1.44	1.90	-0.37	1.80	1.15
	(4)	1.41	2.47	0.40	1.47	1.92	0.54	1.88	1.60
Normal Inverse Gaussian	(3)	-0.56	0.23	0.73	-0.04	0.64	-1.16	0.31	0.02
	(4)	0.57	0.40	0.78	0.25	0.68	1.24	0.39	0.69
Scaled Student-t	(3)	-0.07	-1.32	0.05	-1.67	0.03	-2.59	-1.07	-0.95
	(4)	0.08	1.59	0.52	1.86	0.53	2.60	1.36	1.47
General error distn	(3)	-0.01	0.90	0.82	0.68	1.06	-0.01	0.53	0.57
	(4)	0.07	1.00	0.86	0.70	1.07	0.22	0.62	0.74
DRS	(3)	-0.01	0.06	0.69	0.39	0.67	0.15	0.23	0.31
	(4)	0.04	0.40	0.78	0.50	0.75	0.33	0.35	0.51
Summary	(3)	0.05	0.60	0.65	0.43	0.96	-0.23	0.59	
	(4)	0.58	1.47	1.00	1.22	1.22	1.37	1.21	

Note: (1) Mean kurtosis; (2) SD of kurtosis; (3) Mean error; (4) rmse

Table 2. Examples of density forecasts of financial time series

Authors	Data/Frequency/Horizon	Models	Accuracy Measure
De Gooijer & Zerom (2000)	90 day US T-bill rate Weekly 1 week to 5 weeks ahead	3 Kernel based predictors	KS on uniform cdfs Berkowitz
Schittenkopf, Dorffner & Dockner (2000)	FTSE 100 Index Daily 1 day ahead	Neural Network GARCH GARCH-t	Average log(density)
Weigend & Shi (2000)	S&P 500 Index Daily 1 day ahead	Gaussian Gaussian mixture GARCH(1,1) Hidden Markov Experts Gated Experts	Average log(density) KS
Bauwens, Giot, Grammig & Veredas (2004)	Trade duration, price duration, volume duration taken from intra-day NYSE data	Range of autoregressive conditional duration models	Histograms of cdf
Bao, Lee & Saltoglu (2007)	S& P 500 & NASDAQ Daily 1 day ahead	Normal,Student t, GED, Laplace, Double Weibull, Skewed t, inverse hyperbolic sine, mixture of normals, double gamma, Sargan – all with one of eight GARCH models	KLIC
Egorov, Hong & Li (2006)	US Zero coupon bond yields Monthly Unclear	3 Affine structure Random walk	Hong & Li (2005)
Rapach & Wohar (2006)	FX rates (US vs, UK, Germany, France & Japan) Monthly Up to 24 months ahead	Band-TAR ESTAR	Doornik & Hansen (1994)
Hong, Li & Zhao (2007)	FX rates 30 minute Up to 10 hours	Random walk, GARCH, jump, jump GARCH, ARCD, Regime switching	Hong & Li (2005)

Table 3. Analysis of residuals for 'basic' models for the two return series (Estimation region 8 April 1991 to 31 May 1996)

	Brent		WTI	
Density	Gaussian		Gaussian	
GARCH	Symmetric		Asymmetric $\xi = -0.013$	
Mean Reversion	Yes		Yes	
AR order	0		0	
Mean	0.05		0.04	
SD	1.03		1.04	
Excess Kurtosis	1.44		2.18	
ACF of Standardised Residuals				
Lag	**ACF**	**Ljung Box p value**	**ACF**	**Ljung Box p value**
1	-0.001	0.97	0.005	0.80
2	0.002	0.99	-0.042	0.08
3	0.013	0.92	-0.002	0.17
4	0.025	0.68	0.007	0.26
5	0.009	0.77	0.007	0.37
10	0.017	0.15	-0.002	0.65
20	-0.031	0.40	-0.038	0.33
40	0.003	0.32	0.031	0.12
50	0.044	0.05	0.006	0.05

Table 4. Berkowitz statistics for 299 density forecasts over a range of different horizons for seven density functions with differing assumptions about the variance process and mean reversion.

Density	Variance Process	Mean Reversion	Berkowitz Statistic for Horizon H (in				
			H=10	H=30	H=70	H=130	H=260
Brent							
Gaussian	Constant	No	10.7	10.2	5.7	10.3	23.5
Gaussian	Constant	Yes	17.8	40.2	65.5	123.0	237.2
Gaussian	GARCH	No	1.0	2.0	8.0	11.9	24.1
Gaussian	GARCH	Yes	8.1	21.5	43.7	83.0	171.9
Mixture of 2 Gaussians	GARCH	No	0.9	0.8	2.9	3.8	6.5
Mixture of 2 Gaussians	GARCH	Yes	4.3	11.6	22.0	44.9	97.6
Mixture of 2 Gaussians	GARCH	Adaptive	1.1	1.0	3.6	6.1	8.0
Gaussian - Laplace mixture	GARCH	No	2.8	5.5	6.9	10.2	27.1
Gaussian - Laplace mixture	GARCH	Yes	1.7	3.2	10.0	15.7	44.1
Normal Inverse Gaussian	GARCH	No	1.3	2.0	7.1	10.3	24.4
Normal Inverse Gaussian	GARCH	Yes	5.9	15.8	32.9	62.3	130.1
Scaled Student-t	GARCH	No	1.0	1.0	3.2	4.3	6.9
Scaled Student-t	GARCH	Yes	4.8	10.9	19.3	38.4	84.7
Scaled Student-t	GARCH	Adaptive	0.3	0.4	4.0	6.3	5.7
General error distribution	GARCH	No	1.6	2.1	6.9	11.2	24.8
General error distribution	GARCH	Yes	4.1	8.5	15.7	30.0	71.1
DRS	RNGARCH	No	0.7	1.2	4.1	5.2	7.6
WTI							
Gaussian	Constant	No	1.8	2.2	10.8	15.6	22.4
Gaussian	Constant	Yes	9.0	20.1	51.5	101.4	227.9
Gaussian	GARCH	No	2.4	10.2	30.7	47.1	69.8
Gaussian	GARCH	Yes	9.3	25.7	60.3	103.8	196.5
Gaussian	Asym GARCH	No	1.4	6.8	23.1	32.8	44.6
Gaussian	Asym GARCH	Yes	8.4	23.0	50.5	94.7	196.9
Mixture of 2 Gaussians	GARCH	No	0.2	2.3	9.2	9.8	7.0
Mixture of 2 Gaussians	GARCH	Yes	1.1	3.1	8.2	18.2	52.0
Mixture of 2 Gaussians	GARCH	Adaptive	1.7	9.0	23.3	29.5	26.4
Gaussian - Laplace mixture	GARCH	No	2.8	3.8	10.9	25.3	67.6
Gaussian - Laplace mixture	GARCH	Yes	1.2	3.9	12.2	20.1	48.2
Normal Inverse Gaussian	GARCH	No	1.4	6.4	19.5	29.6	49.4
Normal Inverse Gaussian	GARCH	Yes	3.2	10.0	24.8	43.5	91.7
Scaled Student-t	GARCH	No	0.2	3.0	12.6	13.2	8.0
Scaled Student-t	GARCH	Yes	1.3	3.8	9.1	14.7	35.6
Scaled Student-t	GARCH	Adaptive	0.5	3.5	18.7	27.4	24.8
General error distribution	GARCH	No	1.3	6.8	21.0	29.3	40.8
General error distribution	GARCH	Yes	1.6	6.8	21.9	30.0	40.4
DRS	RNGARCH	No	0.8	4.5	15.1	15.8	8.8

Table 5. Mean reversion estimates for Gaussian density with GARCH evolving over thirteen years (shading indicates significant mean reversion).

	Brent			**WTI**		
GARCH	**Symmetric**			**Asymmetric Median value** $\xi = -0.013$		
Origin (yyyymmdd)	**Long term mean**	**reversion rate**	**p value**	**Long term mean**	**reversion rate**	**p value**
19960415	18.13	0.0078	0.03	19.97	0.0037	0.24
19970414	18.22	0.0079	0.01	19.85	0.0058	0.03
19980413	18.03	0.0080	0.01	19.85	0.0058	0.01
19990412	17.86	0.0070	0.00	18.70	0.0050	0.03
20000410	18.34	0.0065	0.00	19.75	0.0052	0.01
20010409	18.73	0.0051	0.00	19.75	0.0052	0.00
20020408	19.37	0.0044	0.00	19.11	0.0057	0.00
20030407	19.99	0.0036	0.01	19.33	0.0040	0.00
20040405	20.97	0.0028	0.03	20.76	0.0040	0.00
20050404	26.69	0.0010	0.31	28.48	0.0009	0.36
20060403	59.34	0.0003	0.68	46.23	0.0004	0.64
20070402	82.38	0.0002	0.69	34.48	0.0004	0.52
20080331	3.48	-0.0002	0.68	9.12	-0.0003	0.62

Table 6. Parameter estimates for the Mixture of 2 Gaussians, Student's t and DRS for Brent and WTI. Estimation region: 9 April 1991 to 26 November 2008. (Standard errors shown in *italics)*

Mixture of 2 Gaussians

	μ	α_0	α_1	β_1	λ_1	σ_1	σ_2
Brent	0.00060	0.000003	0.046	0.949	0.61	0.71	1.34
	0.00026	*0.000001*	*0.005*	*0.006*	*0.05*	*0.03*	
WTI	0.00049	0.000003	0.038	0.957	0.81	0.77	1.65
	0.00026	*0.000001*	*0.005*	*0.006*	*0.03*	*0.02*	

Student's t

	μ	α_0	α_1	β_1	ν_1
Brent	0.00057	0.000003	0.042	0.954	6.49
	0.00026	*0.000001*	*0.005*	*0.006*	*0.65*
WTI	0.00049	0.000002	0.034	0.963	5.07
	0.00026	*0.000001*	*0.005*	*0.005*	*0.40*

DRS*

	β_0	β_2	β_1	λ	$\bar{\gamma}$	$\bar{\mu}$	δ
Brent	0.0000017	0.045	0.952	0.17	1.44	-0.18	0.07
	0.0000006	*0.005*	*0.006*	*0.07*	*0.18*	*0.06*	*0.02*
WTI	0.0000023	0.038	0.957	0.11	1.86	-0.13	0.05
	0.0000007	*0.005*	*0.006*	*0.03*	*0.22*	*0.05*	*0.01*

*Note: The asymmetric GARCH coefficient *c* in the DRS model was not significant for either series and is set to zero in these estimations.

Table 7. Summary statistics of one to ten day ahead standardised forecast errors compared to kurtosis implied by the estimated density (median value from all estimates) (Forecasts cover 2 April 1996 to 10 December 2008)

			Brent	**WTI**
Mixture of 2 Gaussians	**Errors:**	Mean	0.00	0.00
		SD	1.03	1.05
		excess kurtosis	1.53	2.36
	Density implied	excess kurtosis	1.45	2.57
Student's t	**Errors:**	Mean	0.00	0.00
		SD	1.03	1.03
		excess kurtosis	1.54	2.40
	Density implied	excess kurtosis	3.34	14.57
DRS	**Errors:**	Mean	0.00	0.00
		SD	1.02	1.02
		excess kurtosis	1.52	2.33
	Density implied	excess kurtosis	1.58	2.35

Figure 1. An annotated plot of the crude oil price from 1970 to November 2008

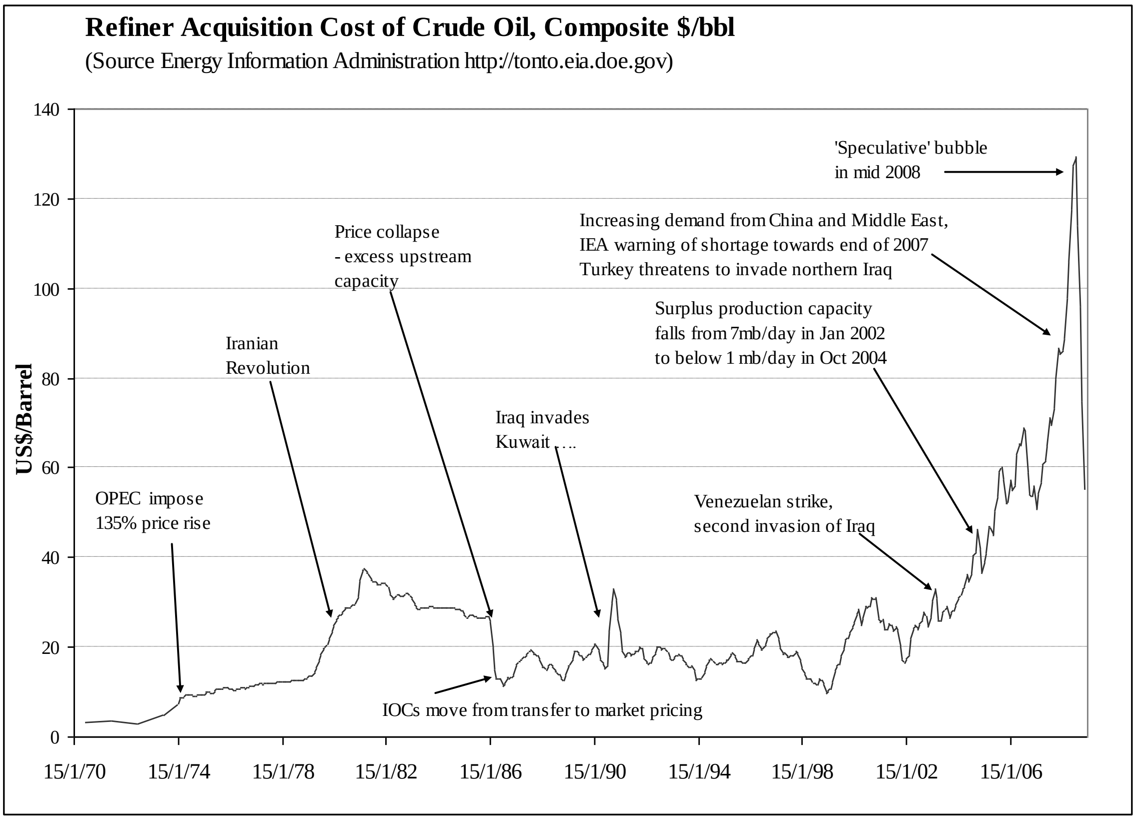
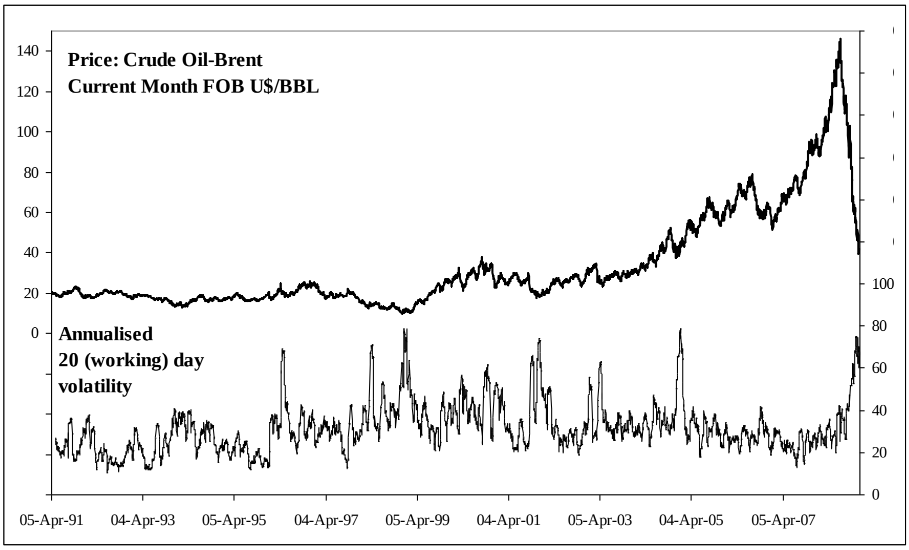


Figure 2. Daily data (April 1991 to September 2008): prices and volatility for both data sets



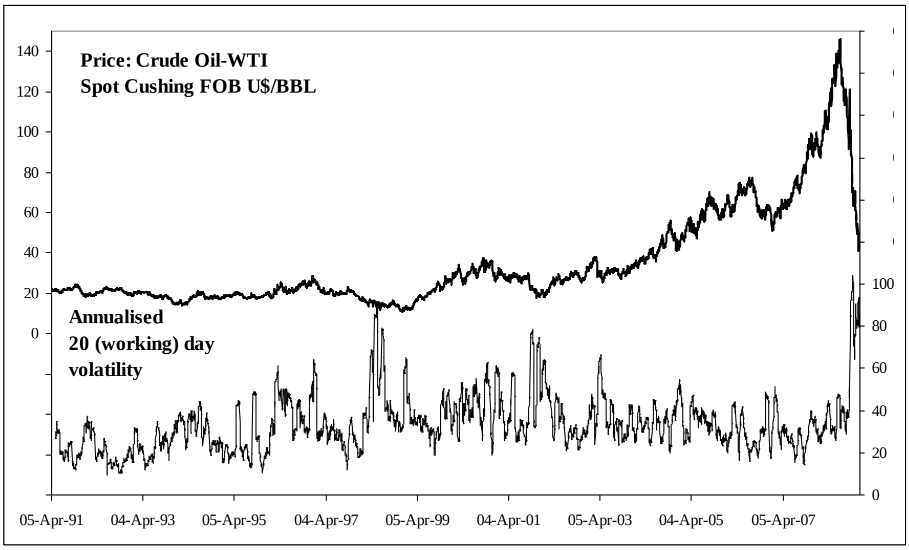


Figure 3. Daily data (April 1991 to September 2008): rolling annual inter-quartile ranges and outliers for both data sets

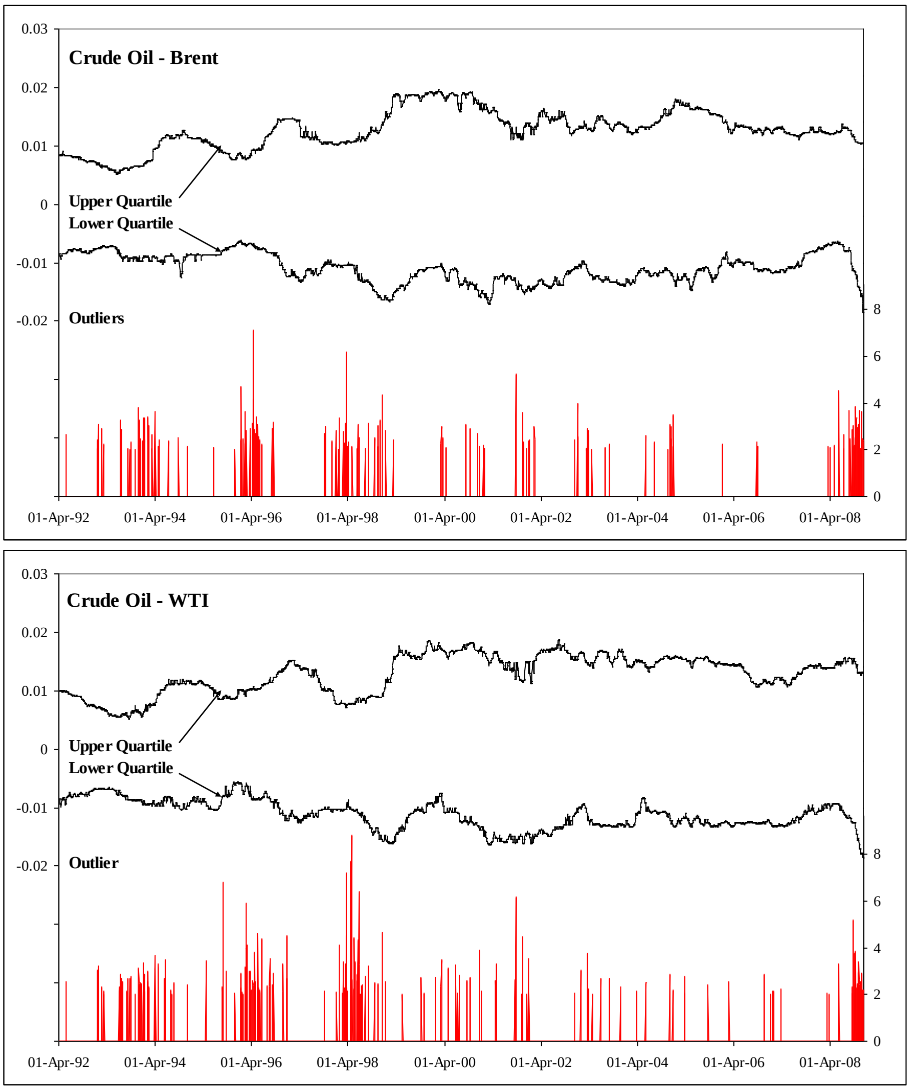


Figure 4. The ln(price) for Brent showing the estimated long term mean using the mixture of two Gaussians with a constant mean and with an adaptive mean.

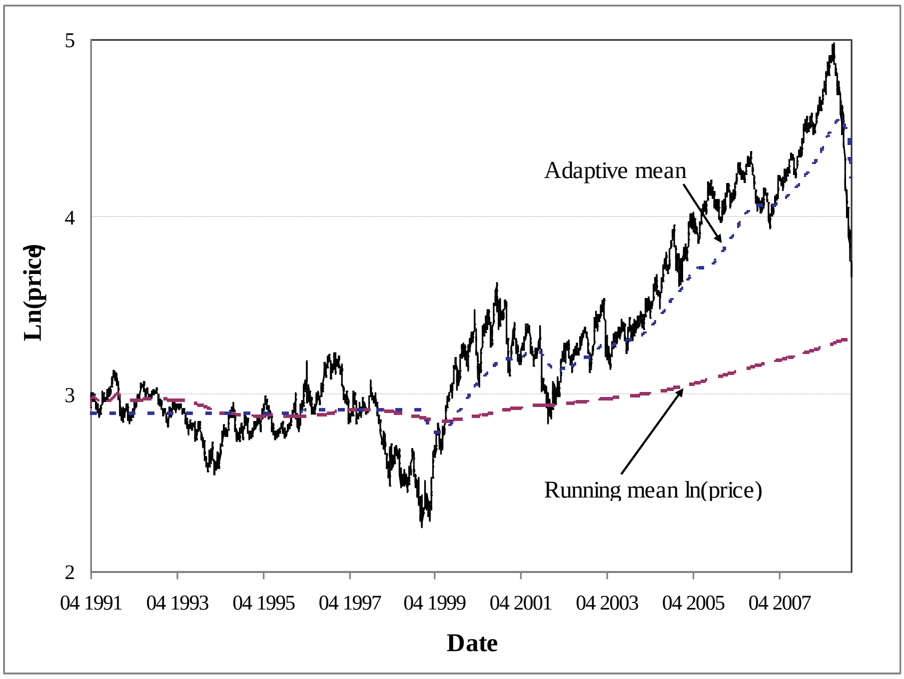


Figure 5. Values of Berkowitz Statistic for density forecasts (using a mixture of two Gaussians) of Brent and WTI prices up to two and a half years ahead (using the maximum number of forecasts available for the most recent data and one and two years previously)

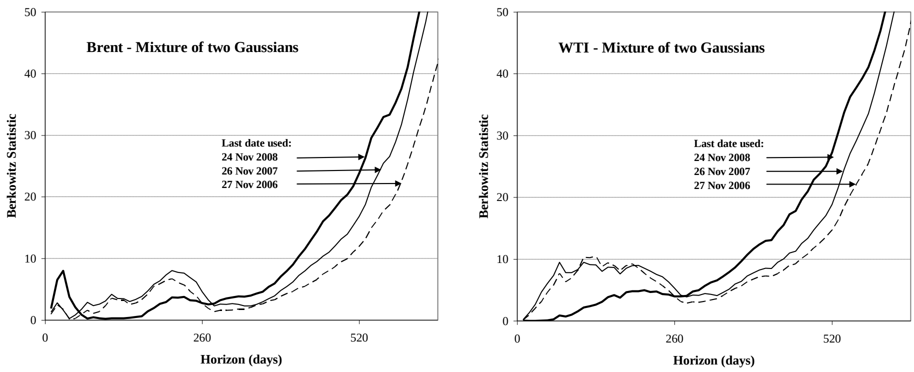


Figure 6. Histograms of empirical cumulative densities of prices for Brent (using a mixture of two Gaussians) for one, two and three years ahead (using the maximum number of available forecasts)

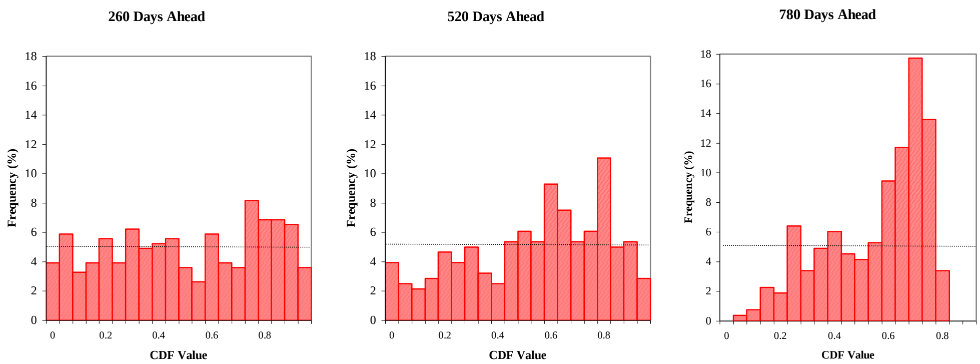


Figure 7. Using an origin of 26th Nov. 2007, the density forecasts of the price of Brent from a mixture of two Gaussians is shown as (left) a time series of percentiles summarising density forecasts until the 24th Nov. 2008 and (right) the predicted density for the price of Brent for 24th Nov. 2008 (showing the same percentiles)

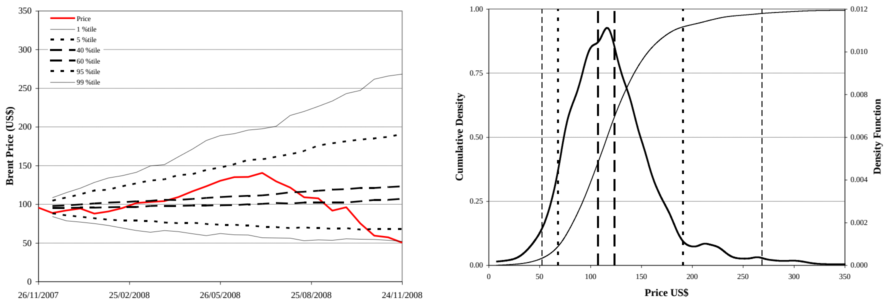


Figure 8. Using an origin of 24th Nov. 2008, the density forecasts of the price of Brent from a mixture of two Gaussians is shown as (left) a time series of percentiles summarising density forecasts until the 23rd Nov. 2009 and (right) the predicted density for the price of Brent for 23rd Nov. 2009 (showing the same percentiles)