ELSEVIER

Energy Economics 24 (2002) 71–95

Energy Economics

www.elsevier.com/locate/eneco

A Markov switching model of the conditional volatility of crude oil futures prices

Wai Mun Fong*, Kim Hock See

Department of Finance and Accounting, National University of Singapore, Kent Ridge Cresent, 119260 Singapore

Abstract

This paper examines the temporal behaviour of volatility of daily returns on crude oil futures using a generalised regime switching model that allows for abrupt changes in mean and variance, GARCH dynamics, basis-driven time-varying transition probabilities and conditional leptokurtosis. This flexible model enables us to capture many complex features of conditional volatility within a relatively parsimonious set-up. We show that regime shifts are clearly present in the data and dominate GARCH effects. Within the high volatility state, a negative basis is more likely to increase regime persistence than a positive basis, a finding which is consistent with previous empirical research on the theory of storage, e.g. Fama and French (1988a,b) and Ng and Pirrong (1994). The volatility regimes identified by our model correlate well with major events affecting supply and demand for oil. Out-of-sample tests indicate that the regime switching model performs noticeably better than non-switching models regardless of evaluation criteria. We conclude that regime switching models provide a useful framework for the financial historian interested in studying factors behind the evolution of volatility and to oil futures traders interested short-term volatility forecasts. © 2002 Elsevier Science B.V. All rights reserved.

Keywords: Crude oil futures; Conditional volatility; GARCH; Markov switching

1. Introduction

Empirical studies indicate that commodity prices can be extremely volatile at times. Webb (1987) describes such occasional outburst of volatility as indications of a 'choppy' market. According to Webb (1987)):

* Corresponding author.

PII: S0140-9883(01)00087-1

> 'a market condition traders frequently encounter in futures markets is the phenomenon of 'choppy market'...where prices oscillate wildly without seeming to stop at intervening prices... Traders fear such markets because of the enhanced danger of uncontrollable losses due to sudden price swings'.

Sudden changes in volatility also have important implications for the pricing of commodity derivatives and the construction of optimal hedge ratios. For example, commodity options will be under priced if the historical unconditional volatility is assumed during periods when there is a switch from a 'low' to 'high' volatility regime. Similarly, highly inaccurate hedge ratios may result when they are computed on the assumption that there are no sudden changes in volatility. Lastly, since the spread between spot and futures prices tend to increase during periods of high volatility, commodity traders will face higher basis risks, which compounds the problem of determining an optimal hedge ratio.

This paper examines the temporal behaviour of volatility of daily returns for crude oil futures and its relation to the basis, i.e. the percentage difference between contemporaneous spot and futures price. Our data shows that oil futures exhibit backwardation 71% of the time. This high frequency of backwardation is consistent with the findings of Litzenberger and Rabinowitz (1995) who explain the phenomenon in terms of option pricing theory. Specifically, oil wells are viewed as call options with exercise price corresponding to the extraction cost. In this context, persistent backwardation is viewed as an inducement for current extraction. Moreover, the more volatile are demand and supply shocks, the higher must the spot price be relative to futures price to overcome producers' preference to leave oil underground. Litzenberger and Rabinowitz (1995) provide empirical evidence based on US data which are consistent with this hypothesis. Specifically, they find that controlling for the level of spot price, there is a significant negative relation between implied volatility of at-the-money-futures call options and the basis. The Litzenberger–Rabinowitz analysis, however, does not preclude the role of shortages in oil stocks in affecting volatility. It is well known that the cost of storing oil above ground is extremely high (Verleger 1994). Thus, oil refiners hold very minimal stocks of oil, accepting the risks of running out of product. According to Verleger (1994), stocks of crude oil held by oil refiners in the US, Europe and Japan typically cover less than 30 days of use. High storage costs severely hinders the ability of oil refiners to cope with situations of sudden high demand, giving rise to inelasticity of oil supply especially in the short-run. Due to inelastic supply, spot prices must rise significantly to equilibrate the market imbalances whenever there is a big increase in demand. Thus, the theory of storage predicts that backwardation is more likely to occur when oil stocks are low than when there are adequate stocks. A dramatic example of backwardation in the crude oil market occurred in 1996–1997 due to depleting oil stocks following the severe Northern Hemisphere winter of 1995, and a surge in oil demand in 1996 by the recovering economies of US, Europe and the Far East (see Petroleum Economist 1990). By the end of 1996, the spot price of crude oil reached US$27, the highest level since 1990. Interestingly, volatility of spot and futures oil prices also rose significantly between 1996 and 1997, reflecting the combined effects of inelastic supply and volatile demand. Based on our analysis, the monthly standard deviation of nearby West Texas Intermediate (WTI) crude oil futures in 1996–1997 was approximately 63% above the level of volatility in 1995. A similar but shorter period of backwardation and surge in volatility occurred in early 1994 which lasted for 6 months. According to industry observers, the backwardation in crude oil reflected significantly higher energy use from Europe and the US (Wall Street Journal, 1995).

To summarise, the options framework relies on demand or supply shocks to explain the increase in the volatility of spot oil prices while the theory of storage places more emphasis on low inventories and volatile demand shocks as the main factors behind big swings in oil prices. That is, according to storage theory, demand shocks by themselves are insufficient to explain increases in volatility without inelastic supply. These two views are, however, not mutually exclusive. For example, during periods when demand shocks are risky, price volatility can persist or increase further if the market goes into backwardation due to a major disruption in supply, as events in 1996–1997 seem to confirm. The purpose of this paper is to attempt to model these complex features of volatility by using a general regime switching model.

Our model builds on the standard GARCH approach by allowing for jumps in the conditional variance between a discrete number of states or regimes. The model is flexible in that all GARCH parameters can switch between regimes. Moreover, the regimes are treated as a latent (unobservable) variable that can be estimated along with the other parameters of the model using maximum likelihood. The transition between volatility states is assumed to be governed by a Markov process. We hypothesise that backwardation will have a more significant effect on the persistence of volatility shocks during periods of high volatility than during periods of low volatility. To capture the possible effects of basis on the persistence of volatility regimes, we allow transition probabilities to be a time-varying function of the basis.

The remainder of this paper is organised as follows. Section 2 introduces the regime switching model and discusses estimation issues. Section 3 describes the data set and presents preliminary descriptive analysis of the data. Estimation results are presented in Section 4. In Section 5, we evaluate the forecasting ability of the regime switching model against two benchmarks: a constant variance and the single regime GARCH (1,1) model. Section 6 concludes the paper.

2. Regime switching models

2.1. Structure of regime switching models

Many financial time series undergo alternating periods of calm and turbulence. In particular, asset prices tend to exhibit volatility clustering in which large (small) price changes tend to be followed by large (small) price changes. This clustering of volatility strongly suggests that conditional volatility of asset returns is time varying.

The most popular approach for modelling conditional volatility is the GARCH family of models as introduced by Engle (1982) and generalised by Bollerslev (1986) and Nelson (1991). GARCH models are appealing because of their simplicity, ease of estimation and empirical success in modelling time-varying volatility in a variety of contexts (Bollerslev et al., 1992). Empirically, a common finding is that GARCH models tend to impute a high degree of persistence to the conditional volatility. This means that shocks to the conditional variance that occurred in the distant past continue to have non-trivial effects currently. The degree of volatility persistence for GARCH models can be easily quantified. Consider the widely used GARCH (1,1) model with h_t as the conditional variance and σ^2 as the unconditional variance. The jth period ahead forecast of the conditional variance is:

$$E(h_{t+j}) = \sigma^2 + (\alpha + \beta)^j (h_t - \sigma^2), \text{ for } j \geq 1 \tag{1}$$

Thus, when $\alpha + \beta$ is close to one, shocks to the conditional variance are highly persistent, and the conditional variance is possibly integrated (Engle and Bollerslev, 1986). Empirical studies typically find $\alpha + \beta$ to be close to one for a variety of assets. For example, French et al. (1987), Chou (1988), Baillie and DeGennaro (1990) and Fong (1997) all report $\alpha + \beta$ to be above 0.9 for weekly stock returns. However, Lamoureux and Lastrapes (1990) point out that such high levels of volatility persistence may be spurious if there are structural breaks or regime shifts in the volatility process. They demonstrate this point by introducing deterministic shifts in the variance and find that this leads to a marked reduction in the degree of volatility persistence compared to that implied by standard GARCH models. This suggests that to obtain more robust estimates of conditional volatility would require a more general class of GARCH models that allows for regime shifts as part of the data generating process.

Regime switching GARCH models were introduced recently by Hamilton and Susmel (1994), Cai (1994) and Gray (1996a,b). There are several common features in these models. First, the conditional volatility process is allowed to switch stochastically between a finite number of regimes. Second, the timing of regime switch is usually assumed to be governed by a first-order Markov process. The transition probability of the Markov process determines the probability that volatility will switch to another regime, and thus the expected duration of each regime. Transition probabilities may be constant or a time-varying function of exogenous variables. Within this framework, several versions of regime switching models have appeared in the literature. Hamilton and Susmel (1994) and Cai (1994) consider regime switching models with ARCH innovations. Gray (1996a) introduced a more general regime switching model that allows for GARCH dynamics. The extension from ARCH to GARCH is made possible by a new algorithm that solves the well-known path dependency problem associated with GARCH models. The algorithm also simplifies the estimation of more complicated models with regime dependent parameters and time-varying transition probabilities. This opens the way for estimating a richer class of stochastic processes for financial assets with complex volatility structures.

We apply a regime switching model similar to Gray's to examine the temporal behaviour of the conditional volatility of crude oil futures. We begin by illustrating the set-up of a basic Gray-type model with two states and GARCH effects only. The model can easily incorporate ARMA dynamics in the mean as well, as we show in the empirical section. The basic model can be written as follows.

$$\begin{aligned} r_{it} &= \mu_{it} + \varepsilon_{it} \\ \varepsilon_{it}|\Omega_{t-1} &\sim N(0,h_{it}) \quad i = 1,2 \text{ states} \end{aligned} \tag{2}$$

where r_{it}, denote returns, and μ_{it} and h_{it} are the conditional mean and conditional variance, respectively, both of which are allowed to switch between two regimes. Regime switching is assumed to be directed by a first-order Markov process with transition probability given by:

$$\left.\begin{aligned} Pr[S_t = 1|S_{t-1} = 1] &= P \\ Pr[S_t = 2|S_{t-1} = 1] &= 1 - P \\ Pr[S_t = 2|S_{t-1} = 2] &= Q \\ Pr[S_t = 1|S_{t-1} = 2] &= 1 - Q \end{aligned}\right\} \tag{3}$$

The above set-up is similar to a mixture-of-distribution model in which the innovations ε_{it} are drawn from a finite number of normal distributions. In the case of regime switching models, the mixing variable is a function of the transition probabilities, P and Q, which determines the probability of the system remaining in the same state. The mixing variable for regime switching models is known as the regime probability, which we denote by $p_{it} = Pr(S_t = i|\Omega_{t-1})$. The regime probability is thus the ex-ante probability of a particular state at time t, conditional on information available at $t - 1$ (Ω_{t-1}) and is a key input for forecasting. For a first-order Markov chain, it is straightforward to show that the regime probability is the weighted average of P and $(1 - Q)$ as shown in Eq. (6). Thus, even with fixed transition probabilities, regime probabilities are time-varying. This is a key feature that distinguishes the Markov switching model from classical mixture models. In Gray's model, the conditional volatility, h_t, is assumed to follow a GARCH(p,q) process

$$h_t = \nu + \sum_{i=1}^{p} \alpha_i \varepsilon_{t-i^2} + \sum_{i=1}^{q} \beta_i h_{t-i} \tag{4}$$

where $\nu > 0$, $\alpha_i \geq 0$ and $\beta_i \geq 0$ to ensure that the conditional variance is positive. All variance and mean parameters are regime-dependent.

An appealing feature of regime switching models is that they allow the joint estimation of regime shifts and GARCH effects. Moreover, the relative significance of regime shifts vs. GARCH effects can be tested, although within a non-standard testing framework (Davies 1977, 1987; Hansen 1992; Garcia 1997). Another interesting feature of regime switching models is that the regimes are not assumed to be observable by the econometrician, but can be identified from the estimation process. Details of the estimation procedure for regime switching models are discussed in Section 2.2.

2.2. Estimation issues

Let f_{it} be the conditional distribution of returns, r_t. By assuming normal innovations, f_{it} can be written as:

$$f_{it} = f(r_t | S_t = i, \Omega_{t-1}) = \frac{1}{\sqrt{2\pi h_{it}}} \exp\left\{ -\frac{1}{2} \frac{(r_t - \mu_{it})^2}{h_{it}} \right\} \tag{5}$$

where $i = 1, 2$ states; h_{it} = variance of i at time t; Ω_{t-1} = information set at time $t-1$; μ_{it} = conditional mean return in state i at time t.

Using Bayesian arguments, it can be shown that the regime probability p_{1t} can be written as a simple first-order non-linear recursive function of the transition probabilities and the conditional distribution (see Gray 1996a, pp. 58–61):

$$\begin{aligned} p_{1t} = {} & P\left[\frac{f_{1t-1} p_{1t-1}}{f_{1t-1} p_{1t-1} + f_{2t-1}(1 - p_{1t-1})} \right] \\ & + (1 - Q)\left[\frac{f_{2t-1}(1 - p_{1t-1})}{f_{1t-1} p_{1t-1} + f_{2t-1}(1 - p_{1t-1})} \right] \end{aligned} \tag{6}$$

where $p_{1t} = Pr(S_t = 1 | \Omega_{t-1})$; $f_{1t} = f(r_t | S_t = 1)$; and $f_{2t} = f(r_t | S_t = 2)$

The log-likelihood function (L) for this model is

$$L = \sum_{t=1}^{T} \log[p_{1t} f_{1t} + (1 - p_{1t}) f_{2t}] \tag{7}$$

The log-likelihood function can be constructed recursively using the following expressions for h_t and ε_t:

$$\begin{aligned} h_t &= E[r_t | \Omega_{t-1}]^2 - [E[r_t | \Omega_{t-1}]]^2 \\ &= p_{1t}(\mu_{1t}^2 + h_{1t}) + (1 - p_{1t})(\mu_{2t}^2 + h_{2t}) - [p_{1t}\mu_{1t} + (1 - p_{1t})\mu_{2t}]^2 \end{aligned} \tag{8}$$

and

$$\varepsilon_t = r_t - E[r_t | \Omega_{t-1}] = r_t - [p_{1t}\mu_{1t} + (1 - p_{1t})\mu_{2t}] \tag{9}$$

At each time step, the conditional variance is obtained by aggregating the conditional variances from the two states based on the regime probabilities. This aggregated conditional variance is then used to compute the conditional variance for the next period, etc. By focusing on the regime probability instead of the transition probability as in Cai (1994) and Hamilton and Susmel (1994), the conditional variance at time t is made to depend only on the current regime, and not the entire sequence of regimes up to time t. This solves the problem of path-dependence for GARCH models, thus making feasible, the estimation of regime switching models with GARCH rather than just ARCH innovations. In recognition of the first-order recursive structure of the regime probability, as shown in Eq. (6), Gray labels this class of models as First-Order Regime Switching (FORS) models.

Given initial values for the regime probabilities, and the conditional mean and variance for each state, the log-likelihood function can be constructed and maximised numerically to obtain parameter estimates of the model. It is common practice to assume that the maximum likelihood estimators are consistent and asymptotically normal. Using arguments similar to those in Bollerslev and Wooldridge (1992), Gray (1996b) proves these properties hold for normal quasi-maximum likelihood estimators of FORS models with fixed transition probabilities. This result also applies to models with time-varying transition probabilities that are a function of the history of the data because in this case, it can be shown that the regime probabilities continue to have a first-order recursive structure.

In this paper, we allow the transition probability of switching between volatility states at time t to be dependent on the lagged basis (b_{t-1}). It is appropriate to use b_{t-1} as the information set since it reflects fundamentals immediately prior to the shock that generates the return at t whereas b_t would include the impact of the shock. Using the lagged basis also preserves the first-order recursive structure of the regime probabilities, ensuring that our maximum likelihood estimators will be consistent and asymptotically normal.

The estimation routine generates two interesting by-products in the form of the regime probability and the smooth probability. Recall that the regime probability at time t is the probability that state i will operate at t, conditional on information available up to $t-1$. The regime probability is a key input in forecasting the conditional variance. The other by-product, is the smooth probability, $Pr(S_t|r_T, r_{T-1}, \ldots, r_0)$, which is the probability of a particular state in operation at time t conditional on all information in the sample. The smooth probability allows the researcher to 'look back' and observe how regimes have evolved over time.

3. Data and descriptive analysis

3.1. Data

The data used in this study are daily returns on the second nearest crude oil futures based on the West Texas Intermediate (WTI) Cushing, Oklahoma contract traded on NYMEX. To proxy for the spot price, we use daily prices for the nearest futures contract. All data is obtained from Bloomberg database. Our sample period is from 2 January 1992 through 31 December 1997 for a total of 1506 observations. Daily returns are computed by taking the difference in logarithm of consecutive days' settlement prices. The following section gives a brief description of the contract specifications of the WTI futures contract.

3.2. Contract specifications

Each WTI crude oil futures contract is equivalent to 1000 US barrels (42 000 gallons) of light, sweet crude oil. The WTI is a monthly contract and is available 30 months into the future.

All futures contract prices are quoted in dollars and cents per barrel. The minimum and maximum price fluctuations are restricted to \$0.01 and \$15 per barrel, respectively. The expiration of crude oil futures is on the third business day prior to the 25th calendar day of the month. To safeguard against market manipulation, there is a position limit of 15 000 contracts for all months combined and not exceeding 1000 contracts in the last 3 days of trading in the spot market or 7500 in any month. Despite such limits, WTI crude oil futures contracts are among the most liquid of all commodity futures contracts.

3.3. Preliminary analysis

Table 1 represents descriptive statistics of the data. The mean daily futures return is small (0.46%) compared to the standard deviation (1.49%). The unconditional distribution of the returns is slightly skewed and consistent with many other financial assets, has fatter tails than the normal distribution.

Panel B of the table presents autocorrelation coefficients for the returns and squared returns for a selection of lags (1, 5 and 10) as well as Ljung Box statistics for lag 5. While the returns exhibit a mild degree of autocorrelation, the squared returns are more highly autocorrelated. This is confirmed by the Ljung-Box statistic, which is significant at less than 1%. The autocorrelation pattern of the squared returns is consistent with volatility clustering where large (small) price changes tend to be followed by large (small) price changes over consecutive days. Volatility clustering in commodity returns have been documented in many empirical studies using standard GARCH models. See e.g., Kroner et al. (1993) and Ng and Pirrong (1994). GARCH models typically imply a high degree of volatility persistence, i.e. volatility shocks which occurred in the distant past continue to have a non-trivial impact on current volatility. However, it may be premature to conclude on the basis of the high degree of volatility persistence that conditional volatility is highly predictable. As Diebold (1986), Lamoureux and Lastrapes (1990) and Hamilton and Susmel (1994) show, volatility persistence may also be due to structural breaks in the volatility process. Therefore, at this stage, it is unclear whether the autocorrelations of the squared returns for crude oil futures reflect genuine volatility predictability or is merely a result of regime shifts. We resolve this issue by nesting GARCH effects within a general regime switching model.

Panel C presents correlations of spot and futures returns and squared returns with the basis. Since the basis can fluctuate with interest rates even if there is no

Table 1
Descriptive statistics

Panel A. Futures returns		
Mean	−0.0046	
Median	0.0000	
Min	−7.1799	
Max	6.6177	
Variance	2.2341	
Skewness	−0.14	
Kurtosis	4.78	
Panel B Auto correlations of futures returns and squared returns		
Lag	Returns	Squared returns
1	0.023	0.023
5	0.020	0.012**
10	0.016	0.031**
Ljung-Box Q(5)	29.36**	76.05**
	(0.00)	(0.00)
Panel C. Correlations of spot and futures returns with basis		
Spot	−0.12	−0.21
Futures	−0.04	−0.12
Panel D. Auto correlations of basis		
Lag	Autocorrelation	
1	0.94**	
5	0.81**	
10	0.76**	
Unit root tests	Test statistic	Value of Test Statistic
ADF test	t_μ	−4.37**
	t_τ	−4.67**
Phillips–Perron test	z_μ	−5.62**
	z_τ	−6.13**

The sample period is 2 January 1992 through 31 December 1997 (1506 observations). $Q_1(5)$ is the Ljung–Box statistic for testing the joint significant of autocorrelations of returns or squared for the first five lags. Under the null hypothesis of zero autocorrelations, the Q-statistic is distributed as a χ^2 variable with 5 d.f. (P-value in parentheses). The Augumented Dickey–Fuller (ADF) test for unit roots for the basis is based on the regression: $b_t = \mu + a_1 t + a_2 b_{t-1} + a_3 \sum_{i=1}^{n} \beta_i b_{t-i} + e_t$ where b_t denotes the interest-adjusted basis on day t. The Phillip–Perron test is based on the regression: bt $= \mu + a_1 t + a_2 b_{t-1} + e_t$, where the innovations, e_t, are allowed to be heterogeneous. The null hypothesis for the two tests is that $a_2 = 0$. The ADF t statistics, t_μ and $t_{\tau-}$ are based on the regression with and without time trend, respectively, as are the Phillip–Perron test statistics, z_μ and z_τ. Ninety-five percent asymptotic critical values for the ADF and Phillip–Perron tests are −2.86 and −3.41, respectively, * and ** denote statistical significance at 10 and 5%, respectively.

changes in demand and supply, we use an interest-adjusted basis (ADB_t) computed as follows:

$$\mathrm{ADB}_t = \frac{F_t - S_t}{S_t} - R_t \tag{10}$$

where F_t and S_t are the futures and spot prices at the close of day t and R_t is the time to expiration yield on 3-month US Treasury bills. Table 1 shows that both spot and futures returns and their squares vary inversely with the basis. In terms of magnitude, the basis is more highly correlated with the spot than with futures returns or squared returns. Thus, the spot price is more sensitive to changes in the basis than the futures price. Furthermore, the smaller the basis, the more volatile are spot returns, and to lesser extent, futures returns. These results are consistent with the theory of storage. See Fama and French 1988a,b; French (1986) and Ng and Pirrong (1994).

Finally, panel D shows that the basis is highly autocorrelated, with a first order autocorrelation at 0.94. Subsequent autocorrelations decay at a slow rate. However, based on the Augmented Dickey Fuller and semiparametric Phillips-Perron test for unit roots, we reject the null hypothesis of a unit root in the basis. Thus, basis follows a stationary but highly persistent stochastic process. The basis persistence, which probably reflects long production lags from the extraction of raw crude oil in the wells to the distribution of refined crude products, has important implications for volatility of spot and futures oil prices. In particular, critical shortages may lead to prolonged periods of significant backwardation as well as volatility for spot and nearby futures prices.

4. Estimation results

Table 2 reports the estimation results for a standard GARCH(1,1) model and a simple Markov switching model without GARCH effects. For both models, the returns errors (ε_t) are assumed to follow a student-t distribution which allows for fatter tails than the Gaussian distribution. The GARCH(1,1)-t model is:

$$r_t = \mu + \varepsilon_t \tag{11}$$

$$\varepsilon_t \sim \text{Student-}t\ (0, \mathrm{FR})$$

$$h_t = \nu + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1} \tag{12}$$

where FR(> 0) is the degree of freedom of the Student-t distribution, and $\alpha_1 \geq 0$ and $\beta_1 \geq 0$, so that the conditional variance, h_t, is always positive.

Consistent with the findings of previous research, the GARCH(1,1) model imputes an extremely high degree of volatility persistence. Since $\alpha_1 + \beta_1 = 0.9957$,

Table 2
Comparison of GARCH(1,1) − t and Markov switching model for crude oil futures

Parameters	GARCH(1,1) − t	Markov switching model
μ_1	0.0083	0.0130
	(0.0188)	(0.0389)
μ_2		0.0021
		(0.0672)
ν_1	0.0135**	1.2300***
	(0.0080)	(0.1182)
ν_2	–	3.2140***
		(0.2425)
α_1	0.0349***	–
	(0.0093)	
β_1	0.9608***	–
	(0.0101)	
P	–	0.9941***
		(0.0035)
Q	–	0.9935***
	(0.0039)	
Fr1	5.7295***	5.6805***
	(0.8303)	(1.2766)
Fr2	–	8.9791***
		(2.9886)
Log-likelihood	−2652.46	−2646.98
LR test	–	10.96**
		(0.08)*

This table reports maximum likelihood estimates of GARCH(1,1)-t and Markov switching models for daily returns on the West Texas Intermediate crude oil futures. The sample period extends from 2 January 1992 to 31 December 1997, (1506 observations). The GARCH(1,1)-t specification is:
$r_t = \mu + \varepsilon_t$; ε_t
$\sim$ Student-t (0, FR_1);
$h_t = \nu_1 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1}$.

The Markov switching specification is:
$r_{it} = \mu_{it} + \varepsilon_{it}$, where $i = 1$; 2 states:
$\varepsilon_{it}|\Omega_{t-1} \sim$ Student-t (0, FR_1).

The transition probabilities of the model are:

$$Pr[S_t = 1|S_{t-1} = 1] = P$$
$$Pr[S_t = 2|S_{t-1} = 1] = 1 - P$$
$$Pr[S_t = 2|S_{t-1} = 2] = Q$$
$$Pr[S_t = 1|S_{t-1} = 2] = 1 - Q$$

Standard errors of parameters are in parentheses. The likelihood ratio (LR) test is computed as follows: -2*(likelihood of H_1−likelihood of H_0), where H_0 is the GARCH(1,1)-t model and H_1 is the Markov switching model with constant transition probabilities. For the Markov switching model, the P-value is computed using Davies's (1977, 1987) adjustment for the problem of nuisance parameters under the null hypothesis of a single regime. *, ** and *** denote statistical significance at 10, 5 and 1%.

on the average, it takes 720 days for volatility shocks to die out ($0.9957^{720} = 0.045$). However, Lamoureux and Lastrapes (1990) provide simulation evidence which shows that the high degree of volatility persistence implied by standard GARCH models may be spurious in the presence of structural breaks in the conditional variance. Hamilton and Susmel (1994) confirm this conjecture using a regime switching model that allows ARCH innovations. Applying their model to weekly US stock index returns, they find strong evidence of regime shifts in the scale of the ARCH process. Moreover, accounting for regime shifts led to a marked reduction in the degree of residual volatility persistence. To deal with the possibility that volatility clustering in crude oil returns is due to structural changes in volatility, we consider a simple regime switching model with no GARCH effects. The model assumes that the volatility process can switch between two states, where ν_1 and ν_2 are the variance in states 1 and 2, respectively. The transition probabilities are assumed to be governed by a first-order, time homogeneous Markov process as follows:

$$\Pi = \begin{bmatrix} P & 1-P \\ 1-Q & Q \end{bmatrix} \tag{13}$$

where $P = Pr[S_t = 1|S_{t-1} = 1]$ and $Q = Pr[S_t = 2|S_{t-1} = 2]$. The second column of Table 2 reports the estimation results for this Markov switching model. There are three interesting observations. First, the volatility in the two states is very different, suggesting two distinct volatility regimes.

Specifically, the variance in state 2 (ν_2) is at least 2 times higher than the variance in state 1 (ν_1). Second, the transition probabilities, P and Q are close to one, indicating that the volatility regimes are highly persistent. Thirdly, the Markov switching model appears to fit the data better than the GARCH model. The log-likelihood of the Markov switching model (-2646.98) is much larger than the GARCH(1,1)-t model (-2652.46). To assess whether the difference in log-likelihood is statistically significant, we compute the standard likelihood ratio statistic, but adjust the P-value of this statistic upward to reflect the problem of nuisance parameters. This problem arises with Markov switching models because under the null hypothesis of a single regime, the states are not identifiable which violates one of the key assumptions that justify the use of likelihood ratio test. To adjust the P-value, we use the method of Davies (1977, 1987) who applies empirical process theory to derive an upper bound for type I error of a modified LR statistic under the null, assuming nuisance parameters are known under the alternative. Suppose M is the P-value from the likelihood ratio test. Davies showed that

$$Pr[LR(q^*)] > M = Pr(\chi^2_d > M) + \frac{VM^{(d-1)/2}e^{-M/2}2^{-d/2}}{\Gamma(d/2)} \tag{14}$$

where $Pr(LR(q^*) > M|H_0)$ is the upper bound critical value, LR is the likelihood ratio statistic, q^* is the vector of transition probabilities ($q^* =$ argmax $LnL(q)|H1$) and d is the number of restrictions under the null hypothesis. Davies also proved a simple analytical formula, $V = 2M^{1/2}$, by assuming that there is a unique global optimum for the likelihood function. As Table 2 shows, the Davies adjustment produces a P-value for the likelihood ratio statistic 0.08. We conclude that the volatility of oil futures is better described by a two-state regime switching model than a single regime GARCH model. This also implies that the high degree of volatility persistence implied by the GARCH model reflects the persistence of volatility regimes rather than true volatility predictability.

We proceed to explore more general Markov switching specifications that allow for time-varying transition probabilities. Before considering details of the model specifications, we discuss two factors that may affect the volatility of futures returns. First, crude oil prices can be expected to be more volatile during winter seasons due to high demand for heating oil (Gray, 1995), a by-product of crude oil which may be exacerbated by freak weather conditions. To allow for seasonality effects, we include a winter dummy variable (D_{1t}) in the conditional variance equation for the periods from October through February. Second, Samuelson (1965) argues that volatility tends to increase when a futures contract is approaching the expiration date, possibly due to traders adjusting their exposures out of the existing contract to the next nearest contract (see Serletis 1992; Gray 1995). Moreover, contracts that are close to maturity will tend to react more to new information than contracts far away from maturity. To account for possible maturity effects, we include a maturity dummy variable (D_{2t}) using a 5-day window period before the expiry date of the nearest contract.

We estimate three versions of regime switching models, all with time-varying transition probabilities and Student-t density. The models are:

- Model 1: RS-t
- Model 2: RSARCH(4)-t
- Model 3: RSGARCH(1,1)-t

The simplest regime-switching model (RS-t) allows the transition probabilities to vary with the lagged basis (b_{t-1}). We use b_{t-1} as the information set in the transition probabilities specification since this reflects market fundamental prior to the shock that generates the return at t, whereas b_t will include the impact of the shock. The transition probabilities can be written as:

$$\begin{aligned} P_t &= \phi(c_1 + d_1 b_{t-1}) \\ Q_t &= \phi(c_2 + d_2 b_{t-1}) \end{aligned} \tag{15}$$

where b_{t-1} denotes lagged basis; $\phi(\cdot)$ is the cumulative normal distribution function to ensure that the transition probabilities, P_t and Q_t, lie within the unit interval. The model proper is:

$$\begin{aligned} r_{it} &= \mu_i + \varepsilon_{it} \\ \varepsilon_{it} &\sim Student\text{-}t(0, \mathrm{FR}_i) \\ h_{it} &= \nu_i + \lambda_{i1} D_{1t} + \lambda_{i2} D_{2t} \end{aligned} \tag{16}$$

where i = state 1 or state 2; μ_{it} is the conditional mean return; ε_{it} denote the returns innovations; ν_i (> 0) is the unconditional variance; and λ_{i1} and λ_{i2} are coefficients for the winter dummy (D_{1t}) and maturity dummy (D_{2t}), respectively.

Model 2 (RSARCH-t) combines regime switching volatility with ARCH effects within each regime. It extends the switching ARCH model of Hamilton and Susmel (1994) by letting all volatility parameters to switch across regimes, whereas the Hamilton–Susmel model allows only the scale of the ARCH process to switch. Based on an examination of the autocorrelations of squared returns, we decide that an ARCH(4) specification for the conditional variance would suffice. The specification for the conditional variance for the RSARCH(4)-t model is thus:

$$h_{it} = \nu_i + \lambda_{i1} D_{1t} + \lambda_{i2} D_{2t} + \sum_{p=1}^{4} \alpha_{ip} \varepsilon_{t-p}^2 \tag{17}$$

where $\alpha_{ip} \geq 0$ and $\nu_i > 0$.

Model 3 (RSGARCH-t) extends the previous specification by incorporating GARCH effects within each regime. In our estimation, we employ the widely used GARCH(1,1) specification. This model is similar to Gray's (1996b) in his analysis of US Treasury bill rates.

The conditional variance for this model can be written as:

$$h_{it} = \nu_i + \lambda_{i1} D_{1t} + \lambda_{i2} D_{2t} + \alpha_{i1} \varepsilon_{t-1}^2 + \beta_{i1} h_{t-1} \tag{18}$$

where $\alpha_{i1} \geq 0$, $\beta_{i1} \geq 0$ and $\nu_i > 0$.

The estimation results for the three models are reported in Table 4. The RS-t model has a larger log-likelihood of -2638.43 compared to -2646.98 for the simple Markov switching model with constant transition probabilities (refer to Table 2). Based on the likelihood ratio test, we reject the null hypothesis of constant transition probabilities in favour of the RS-t specification at the 1% level of significance. This implies that the assumption of time-varying transition probabilities model results in a model that fits the data better than a fixed transition probabilities model.[1]

Adding ARCH effects improves the ability of Model 1 to capture the data. The log-likelihood of the RSARCH(4)-t model. The log-likelihood improves from -2638.43 to -2627.57.[2] Based on the RS-t model as the null, the likelihood ratio statistic is 21.72 which allows us to reject the RS-t in favour of the RSARCH(4) model at the 1% level of significance.

[1] To ensure a more direct comparison between the fixed transition probability and time-varying transition probability models, we estimate a time-varying transition probability model without dummy variables in the conditional variance equation. Our results show that the log-likelihood of the new model improves to -2643.85. Based on the likelihood ratio test, we again reject the null hypothesis of fixed transition model in favour of the time-varying transition probability model (Model 2 above) at less than 5% level of significance.

[2] Since the first, second and third order of ARCH parameters persistently fall onto the zero boundary restriction, we exclude them in the conditional variance equation for proper convergence.

The final model combines regime switching with GARCH(1,1) conditional variance within regimes. The RSGARCH(1,1)-t model is the most general of the three models. However, the log-likelihood of this model (-2634.69) compares unfavourably with that of the RSARCH(4) model (-2627.57). Moreover, after accounting for regime shifts, only β_{11} is statistically significant in the conditional volatility compared to the standard GARCH(1,1)-t model. This finding supports the conjecture by Diebold (1986) and Lamoureux and Lastrapes (1990) that GARCH effects could be an artifact of the data caused by structural changes in the volatility process. We conclude that of the three models considered thus far, the

Table 3
Markov switching models with time varying transition probabilities

Parameters	Model 1 RS-*t*	Model 2 RSARCH(4)-*t*	Model 3 RSGARCH(1,1)-*t*
μ_1	0.0083 (0.0389)	0.0073 (0.0384)	0.0034 (0.0369)
μ_2	0.0252 (0.0663)	0.0349 (0.0676)	0.0415 (0.0727)
ν_1	1.0922*** (0.1251)	1.0443*** (0.1276)	0.4428** (0.2923)
ν_2	2.7469*** (0.2695)	2.6787*** (0.2867)	2.7239*** (0.9769)
c_1	2.4299*** (0.2212)	2.4358*** (0.2243)	2.4688*** (0.2256)
c_2	2.3157*** (0.2217)	2.3020*** (0.2234)	2.2675*** (0.2576)
d_1	−0.1063 (0.1583)	−0.1021 (0.1626)	−0.0880 (0.1491)
d_2	−0.2732** (0.1521)	−0.2747** (0.1523)	−0.2461** (0.1448)
λ_{11}	0.2256 (0.1913)	0.2210 (0.1887)	0.1232 (0.1246)
λ_{21}	0.5816* (0.4360)	0.6177* (0.4462)	0.6293* (0.4757)
λ_{12}	0.2845* (0.2193)	0.3121* (0.2220)	0.2179* (0.1458)
λ_{22}	1.2407** (0.5656)	1.1624** (0.5777)	1.1921** (0.6609)
α_{11}	–	–	0.0000@
α_{21}	–	–	0.0000@
α_{14}	–	0.0449 (0.0500)	–
α_{24}	–	0.0233 (0.0367)	–
β_{11}	–	–	0.5073** (0.2516)
β_{21}	–	–	0.0246 (0.3349)

Table 3 (*Continued*)

Parameters	Model 1 RS-*t*	Model 2 RSARCH(4)-*t*	Model 3 RSGARCH(1,1)-*t*
Fr1	5.7602***	5.7595***	5.7320***
	(1.2813)	(1.2871)	(1.1189)
Fr2	10.1516***	10.3673***	10.3679***
	(3.7590)	(3.9562)	(4.3512)
Log-likekihood	−2638.43	−2627.57	−2634.69
LR1 test	17.10***	–	–
LR2 test	–	21.72***	–

This table reports maximum likelihood estimates of three Markov switching models: RS-t, RSARCH(4)-t and RSGARCH(1,1)-t for daily returns on the West Texas Intermediate crude oil futures contract. The RS-t specification is:

$r_{it} = \mu_{it} + \varepsilon_{it}$where i = 1,2 states

$\varepsilon_{it}|\Omega_{t-1} \sim$ Student-t(0, FR_i)

$h_{it} = \nu_i + \lambda_{i1}D_{1t} + \lambda_{i2}D_{2t}$

The RSARCH(4)-t specification is:

$r_{it} = \mu_{it} + \varepsilon_{it}$wherei = 1,2 states

$\varepsilon_{it}|\Omega_{t-1} \sim$ Student-t(0, FR_i)

$h_{it} = \nu_i + \lambda_{i1}D_{1t} + \lambda_{i2}D_{2t} + \alpha_{i4}\varepsilon^2_{(t-4)}$

The RSGARCH(1,1)-*t* specification is:

$r_{it} = \mu_{it} + \varepsilon_{it}$where i = 1,2 states

$\varepsilon_{it}|\Omega_{t-1} \sim$ Student-t(0, FR_i)

$h_{it} = \nu_i + \lambda_{i1}D_{1t} + \lambda_{i2}D_{2t} + \alpha_{i1}\varepsilon^2_{(t-1)} + \beta_{i1}h_{(t-1)}$

The transition probabilities of the model is:

$$Pr[S_t = 1|S_{t-1} = 1] = P_t$$
$$Pr[S_t = 2|S_{t-1} = 1] = 1 - P_t$$
$$Pr[S_t = 2|S_{t-1} = 2] = Q_t$$
$$Pr[S_t = 1|S_{t-1} = 2] = 1 - Q_t$$

where $P_t = \phi(c_1 + d_1 b_{t-1})$ and $Q_t = \phi(c_2 + d_2 b_{t-1})$

Standard errors of the parameters are reported in the parentheses. λ_{11} and λ_{21} are coefficients of dummy variable for the winter effect (D_{1t}), and λ_{12} and λ_{22} are coefficients of dummy variable for the maturity effect (D_{2t}). The likelihood ratio test (LR1) is computed as follows: -2*(likelihood of H_1 − likelihood of H_0), where H_0 is the Markov switching model with constant transitions and H_1 is the RS-t model with time-varying transitions. For the LR2 test, H_0 is the RS-t model with time-varying transitions and H_1 is the RSARCH(4)-t model. The sample period is from 2 January 1992 to 31 December 1997, a total of 1506 observations. *, ** and *** denote statistical significance at 10, 5 and 1%, respectively. @ denotes that the parameter estimate fell to the zero boundary.

RSARCH(4)-t model describes the data best. Table 3 represents details of the estimates for the RSARCH(4)-t model.

We see that the unconditional mean returns (μ_{it}) are not significantly different from zero in each regime. Nevertheless, in terms of signs, they are consistent with a positive relationship between expected returns and volatility. The unconditional volatility parameter (ν_2) in the high-variance state is at least twice the parameter in the low-variance state (ν_1). This underscores a need for a model that can account for separate volatility regimes. Only one coefficient in the transition probability equation (d_2) is statistically significant at the 5% level. Since d_2 is negative, this indicates that when returns are in the high-volatility state, shortages will increase the persistence of the high volatility regime. This is consistent with the asymmetry of commodity price reactions to demand shocks as pointed out by Williams and Wright (1991) and Verleger (1994). Likewise, the coefficient for the winter dummy is significant only in the high-volatility state, implying that oil prices are especially sensitive to changes in demand during the winter season when stocks are low. The maturity effect is significant in both the low and high volatility states. This may be explained by the tendency for traders to adjust their positions ahead of expiry of the contract as pointed out by Serletis (1992) and Gray (1995). Finally, within each regime, the residual ARCH effects are small and insignificant, indicating that the volatility process is better described by a regime switching model than a standard ARCH model. The importance of incorporating regime shifts is apparent if we compare the log-likelihood of the RSARCH(4)-t model (-2627.57) with that of a standard ARCH(4)-t model (-2675.30). The likelihood ratio statistic with the ARCH(4)-t model as the null is 95.46 which is significant at any conventional levels. Although not a formal test, the extreme low p-value gives us some degree of confidence to infer that regime shifts are present in the data. Table 4 presents diagnostic statistics of all the models discussed so far.

We now examine the regime and smoothed probabilities generated by the RSARCH(4)-t model to trace how volatility has evolved over the sample period. Fig. 1 displays the regime probabilities and Fig. 2 displays the smoothed probabilities, with the basis superimposed on each plot. The regime probability is of interest in forecasting, while the smoothed probability enables a researcher to 'look back' to determine when a particular regime has emerged. To classify regimes, we follow the Hamilton (1989) scheme where an observation is assigned to state if the probability of that state is higher than 0.5.

Two periods were clearly identified as being in the high-volatility state. The first period was from 26 October 1993 to 26 September 1994, a period of 230 days. This period should be seen in the context of 1994, which was a very volatile year for energy products (Wall Street Journal, January 3 1995). The unconditional standard deviation of daily crude oil returns for 1994 averaged 1.62% compared with 1.29% for the years 1992 and 1993. Oil shortages were a key reason for the jump in volatility. The oil market experienced several oil supply disruptions caused by a civil war in Yemen, civil disturbances in Nigeria and the closure of a major North Sea production platform.[3] Initially, oil stocks were thought to be extremely high due to oversupply in 1993.[4] However, from the second quarter of 1994, strengthen-

[3] 'Crude oil highest in 7 months crude oil hit \$19, drop back slightly' (The Dallas Morning News, 21 May 1994, p. 1A).

[4] 'Crude oil prices decline sharply amid oversupply' (The Wall Street Journal Europe, 7 December 1993, p. 21.

Table 4
Model diagnostic statistics

Model	GARCH(1,1)-t	Markov switching	RS-t	RSARCH(4)-t	RSGARCH(1,1)-t
Log-likelihood (L)	−2652.46	−2646.98	−2638.43	−2627.57	−2634.69
Parameters (K)	5	8	10	12	14
AIC test	−2657.46	−2654.98	−2648.43	−2639.57	−2648.69
SBC test	−2670.75	−2676.25	−2675.02	−2671.47	−2685.91
Standardised Residuals:					
Skewness	−0.18	−0.08	−0.10	−0.14	−0.12
Kurtosis	4.59	4.46	4.45	4.43	4.52
JB test	167.28***	134.69***	134.78***	134.08***	149.35***
$Q_1(5)$	6.56 (0.26)	7.49 (0.19)	6.60 (0.25)	6.05 (0.30)	6.32 (0.28)
$Q_1(20)$	27.45 (0.12)	30.42* (0.06)	27.31 (0.13)	27.35 (0.13)	26.46 (0.15)
$Q_2(5)$	4.51 (0.48)	8.63 (0.13)	9.92* (0.08)	3.84 (0.57)	9.46* (0.09)
$Q_2(20)$	12.75 (0.89)	18.05 (0.58)	17.24 (0.64)	11.07 (0.94)	16.71 (0.67)

The table presents diagnostic statistics for a single regime GARCH model and four regime switching (RS) models. The count of the number of parameters (K) attributed to the ARCH/GARCH specifications does not include the transition probabilities. AIC is Akaike's Information Criterion and SBC is the Schwartz Bayesian Criterion for model adequacy. AIC was calculated as $(L - K)$ for K the number of parameters in the model. SBC was calculated as $[L - (K/2)^* \ln(T)]$ for $T = 1506$. JB is the Jarque–Bera test of normality. $Q_i(q)$ is the Ljung–Box statistic for the joint significant of autocorrelations of standardised residuals for the first q lags. $Q_2(q)$ is the Ljung–Box statistic for the joint significant of autocorrelations of squared standardised residuals for the first q lags. Under the null hypothesis of zero autocorrelations, each of the Q-statistic is distributed as a χ^2 variable with q degrees of freedom (P-value in parentheses). The sample period is from 2 January 1992 to 31 December 1997 (1506 observations). ***, ** and * denote statistical significance at 1, 5 and 10%, respectively.

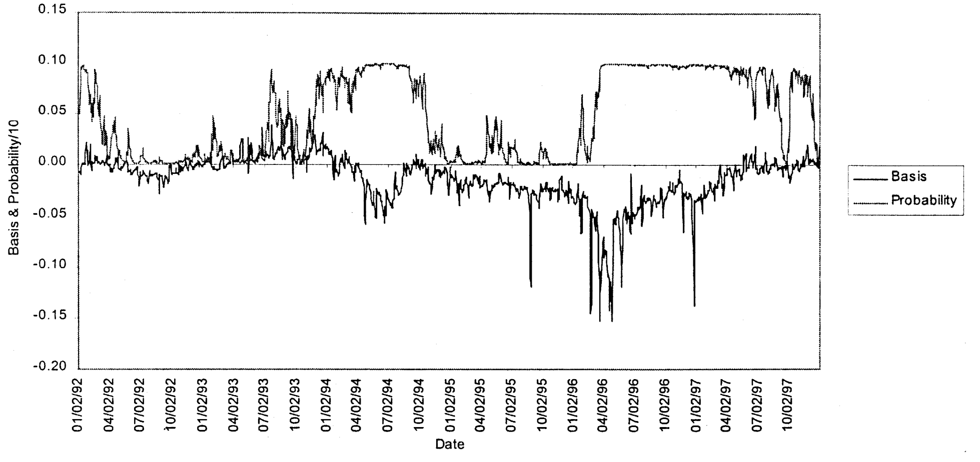

Fig. 1. Basis and regime probabilities of high-volatility state for crude oil futures returns. *Note.* This figure plots the interest-adjusted basis and regime probabilities of the second nearest WTI crude oil futures return. Regime probabilities are generated by a RSARCH(4)-t model. The sample period is from 2 January 1992 to 31 December 1997 (1506 observations).

ing global demand for oil products especially by the U.S. and Europe quickly led to tightness in supply. The possibility of an extended period of shortage was even

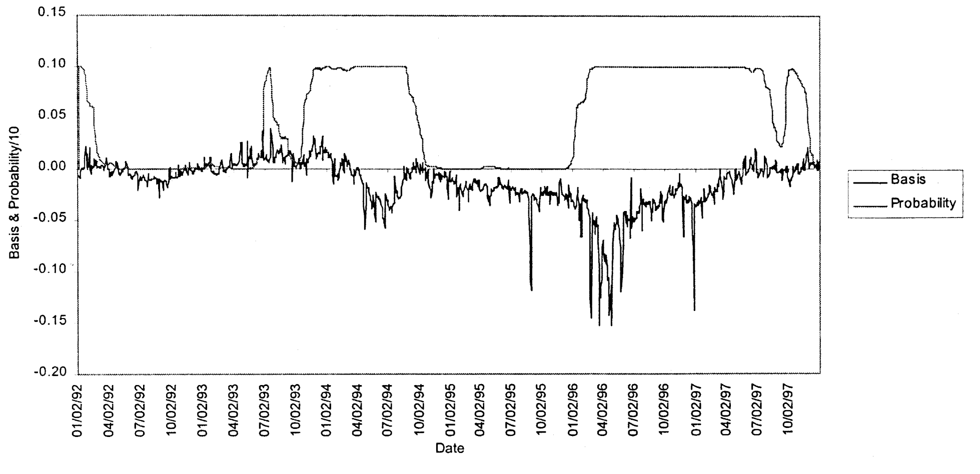

Fig. 2. Basis and smoothed probability of high-volatility state for crude oil futures returns. *Note.* This figure plots the interest-adjusted basis and smoothed probabilities of the second nearest WTI crude oil futures returns. Smoothed probabilities are generated by a RSARCH(4)-t model. The sample period is from 2 January 1992 to 31 December 1997 (1506 observations).

acknowledged by the International Energy Agency (IEA) in early September 1994.[5] Predictably, the prospect of imminent shortages led to a sharp increase in the price of spot crude oil from $14.57 in January to $17.62 in October. The reaction of oil prices is consistent with the predictions of the theory of storage. Figs. 1 and 2 show that the expected effects of backwardation on volatility were largely captured by the regime switching model.

The second high-volatility period lasted 402 days from 11 January 1996 to 14 November 1997. This period again coincided with an exceptional state of backwardation in the oil market during which the basis was persistently negative. The average daily basis over this period was -2.92%. In fact, in 1996 alone, the basis averaged -4.92%. A combination of severe shortage and rising demand explains why the crude oil market was in a state of persistent backwardation during this period. First, an unusually cold 1995/1996 winter spell in the Northern Hemisphere eroded oil stockpiles among major oil producers. Oil inventories remained low throughout 1996 as the Organization of Petroleum Exporting Countries (OPEC) continued to stick with the production ceiling of 24.5 million barrels per day despite increasing oil demand from industrialised economies.[6] The increasing adoption of 'just-in-time' stocking by refiners to keep inventories lean to cut down on carrying costs contributed further to the demand-supply imbalance.[7] In addition, there were concerns about the outcome of oil sales talks between the United Nations and Iraq throughout 1996.[8] Interestingly, oil prices continued to rise while Iraq was getting ready to return to the crude oil exporting business after the approval of the oil-for-food accord in December 1996.

After a brief respite of 1 month, volatility returned to the high-variance state on 23 September 1997 and continued in that state until 25 November 1997, a period of 46 days. This increase in volatility can be traced to political tensions in the Middle East in the last quarter of 1997 as Iraq warned US citizens and aircraft serving with the United Nations arms inspection teams. Oil prices rose on speculation about possible supply disruption when UN condemned Iraq's threats to expel the Americans, but eased in late November 1997 following news that OPEC will increase its

[5] The International Energy Agency raised the estimates for world oil demand in the fourth-quarter by 100 000 barrels, to 69.8 million barrels a day. At this rate, demand was expected to outstrip supply by roughly 3 million barrels a day. See 'Crude Oil Prices Rebound As Demand Looks Better' (The New York Times, 7 September 1994, p. 18).

[6] In August, IEA raised its 1996 demand forecast by 100 000 barrels to 71.8 million barrels a day, up 2.6% from 1995. Meanwhile, stock piles for countries in the Organisation for Economic Co-operation and Development were 98 million barrels below those in 1995 by the end of June. See 'Oil futures hit highest level since April' (The Fort Worth Star-Telegram, 20 August 1996, p. 11).

[7] 'Cold winter spurs rise in demand and prices' (Petroleum Economist, 24 December 1996).

[8] The talks started in May 1995 and centred around United Nations (UN) resolution 986 which allowed Iraq to sell US$l billion worth of oil every 90 days for 180 days on a renewable basis. In effect, an accord on resolution 986 would enable Iraq to supply approximately 700 000 barrels oil daily, an output equivalent to some 10% of leading producer Saudi Arabia. After many rounds of negotiations, the accord was implemented in early December 1996. See 'Review of commodities markets' (The Wall Street Journal, 2 January 1997, p. 34).

Table 5
Chronology of events

Date of high-volatility state identified by RSARCH(4)-t model	No. of days in high-volatility state	Events
26 October 1993–26 September 1994	230	Civil war in Yemen (DMN, 21 May 1994). Civil disturbance in Nigeria (DMN, 21 May 1994) Closure of a major North Sea production platform disrupted daily world oil production (DMN, 21 May 1994). Sudden rise in global demand for crude oil led by US (NYT, 7 September 1994).
11 January 1996–14 August 1997	402	Unusually cold 1995/1996 winter in the Northern Hemisphere (FWS, 20 August 1996). Increasing oil demand from growing world economy (FWS, 20 August 1996). Increasing just-in-time stocking by oil refiners (PE, 24 December 1996). OPEC sticks to its production ceiling of 24.5 million barrels per day despite of increasing pressure to meet oil demand (FWS, 20 August 1996). Oil talks between UN and Iraq for approval of the oil-for-food accord and entry to crude oil market (WSJ, 2 January 1997).
23 September 1997–25 November 1997	46	Build up of tension in Middle East after Iraq refused to allow US to serve with the UN arms inspection teams (WSJ, 31 October 1997). OPEC production ceiling raised to 27.5 million barrels per day (NYT, 29 November 1997).

Sources: DMN, The Dallas Morning News (DMN); FWS, The Fort Worth Star-Telegram (FWS); PE, Petroleum Economist; NYT, The New York Times; WSJ, The Wall Street Journal (WSJ).

production ceiling to 27.5 million barrels per day in respond to the tight supply conditions in the energy market.[9] News of the OPEC increase was immediately reflected in the basis, which reverted back to contango by the end of 1997. Throughout 1996–1997 when the oil market was mainly in a state of severe backwardation, oil prices remained extremely volatile. Specifically, the monthly standard deviation of nearest crude oil futures returns was 63% above the standard deviation in 1995. Table 5 summarises the major events that occurred during the high volatility periods. Figs. 1 and 2 show that our regime switching is able to capture the dramatic changes in volatility profile with remarkable accuracy.

[9] 'OPEC still seeking accord on '98 output' (The New York Times, 29 November 1997, p. 4).

5. Forecasting performance

While our regime switching model appears to fit the data well, it is natural to ask whether the model does a good job in volatility predictions as compared to non-regime switching models. Since regime switching models often involve many parameters, over-parameterisation is a concern. In other words, even though regime switching may be a systematic feature of the data, regime switching models might under-perform non-regime switching models in out-of-sample forecasts because they depend on too many parameters. We address the issue of over-parameterisation by evaluating the out-of-sample forecasting performance of the RSARCH(4)-t model against two benchmarks: constant variance and GARCH(1,1)-t models. To ensure that our evaluations are robust, we use three measures of forecast accuracy and three test periods with different volatility profiles. The three measures of forecast accuracy are: mean square errors (MSE); mean absolute errors (MAE); and R^2. The MSE and MAE are computed as follows:

$$\mathrm{MSE} = \frac{\Sigma(\varepsilon_t^2 - h_t)^2}{n} \tag{19}$$

$$\mathrm{MAE} = \frac{\Sigma|\varepsilon_t^2 - h_t|}{n} \tag{20}$$

In addition, R^2 is computed to measure the goodness-of-fit of the forecasts. The higher the R^2, the more highly correlated is the forecast with the actual volatility. In the forecasting tests, we emphasise models with the highest R^2 because they enable us to obtain direct conclusions about a particular forecast rather than some linear transformations of that forecast (see Gray, 1996b). R^2 is defined as:

$$R^2 = 1 - \frac{\Sigma(\varepsilon_t^2 - h_t)^2}{\Sigma\varepsilon_t^4} \tag{21}$$

where ε_t^2 is the actual volatility and h_t is the forecast volatility at time t.

The forecasting experiment is set up as follows. First, an estimation period is chosen to obtain parameter estimates for each model. Second, based on the parameters estimated, we project a time series of one-step ahead forecasts of the conditional volatility over the test periods. Three test periods are used in the experiment.

- October 1993 to December 1997;
- January 1996 to December 1997; and
- March 1997 to December 1997.

Each test period captures a different volatility profile in the data. In the first test, the estimation period starts before the dramatic increase in volatility in 1994.

Table 6
Out-of-sample forecast performance of Markov switching model

Out-of-sample periods	Statistic	Constant variance	GARCH(1,1)-t	RSARCH(4)-t
October 1993–	MSE	21.67	21.62	21.56
December 1997	MAE	2.80	2.60	2.59
(1050)	R^2	0.23	0.23	0.24
January 1996–	MSE	25.24	25.70	24.72
December 1997	MAE	3.38	3.36	3.17
(500)	R^2	0.28	0.27	0.29
March 1997–	MSE	13.88	14.32	13.41
December 1997	MAE	2.63	2.80	2.61
(200)	R^2	0.29	0.27	0.30

The table evaluates the forecasting performance of the RSARCH(4)-t model against a constant variance model and a GARCH(1,1)-t model. The evaluation criteria are: mean squared errors (MSE), mean absolute errors (MAE) and R^2. Parameters are estimated over the in-sample period, and held fixed for testing over three out-of-sample periods as indicated in the table. The entire sample period extends from 2 January 1992 to 31 December 1997, a total of 1506 observations.

This out-of-sample forecast test will evaluate models' ability to capture unobserved volatility swings in 1994 and 1996. The second test spans the most volatile period (1996), when the energy market was in a state of strong backwardation. The third test covers the year 1997 when volatility subsided.

Results of the forecasting experiment are shown in Table 6. The regime switching model outperformed the constant variance and GARCH models for all three test periods. The RSARCH(4)-t model also performs consistently well in terms of all three forecasting criteria, having lower MSE and MAE and higher R^2 than the other two models. We conclude the RSARCH(4)-t model is not over-parameterised. More generally, our results suggest that regime switching models is a promising tool for delivering more accurate forecasts of short-term volatility than standard GARCH models.

6. Conclusion

This paper examines the temporal behaviour of volatility of daily returns on crude oil futures using a generalised regime switching model. Our model allows for abrupt changes in mean and variance, GARCH dynamics, basis-driven time-varying transition probabilities and conditional leptokurtosis. This flexible model enables us to capture many complex features of conditional volatility within a relatively parsimonious set-up. The results clearly show that regime shifts are present in the data and dominate GARCH effects. Ignoring regime switches in volatility may, therefore, give rise to the false impression that the volatility of oil futures returns is highly predictable. We also find that conditional on the high volatility state, the more negative the basis, the more persistent is the regime, a finding which is consistent with previous empirical research, e.g. Fama and French (1988a,b) and Ng and Pirrong (1994). The volatility regimes identified by our model appear to correlate well with major events affecting supply and demand for oil. In particular, periods of severe shortages and backwardation correspond to periods of extreme volatility in oil futures prices. Although our regime switching model involves many parameters, forecast performance was not compromised as various out-of-sample tests indicate that the regime switching model performed noticeably better than non-switching models regardless of evaluation criteria. We conclude that regime switching models are useful to both the financial historian interested in studying the factors behind the evolution of volatility and to oil futures traders interested in using the model to extract short-term forecasts of conditional volatility. An interesting direction for future research is to explore the relative performance of regime switching models in dynamic hedging compared with strategies using a constant covariance or GARCH-covariance type models (see Gagnon and Lypny, 1997).

Acknowledgements

An earlier version of this paper was presented under the title of 'An Econometric Analysis of Commodity Index Futures with Markov Switching Variance' at the First Conference of the Portuguese Finance Network in June 2000. We would like to thank Professors Lilian Ng, Robert Pindyck and Bruce Hansen for their comments, suggestions and technical assistance. All errors remain ours.

References

Baillie, R.T., DeGennaro, R.P., 1990. Stock returns and volatility. J. Finan. Quant. Anal. 25, 203–215.

Bollerslev, T., 1986. Generalised autoregressive conditional heteroskedasticity. J. Econ. 31, 307–327.

Bollerslev, T., Chou, R.Y., Kroner, K.F., 1992. ARCH modelling in finance: a review of the theory and empirical evidence. J. Econ. 52, 5–59.

Bollerslev, T., Wooldridge, J.M., 1992. Quasi-maximum likelihood estimation and inference in dynamic models with time-varying covariances. Econ. Rev. 11, 143–172.

Cai, J., 1994. A Markov model of switching-regime ARCH. J. Bus. Econ. Stat. 12, 309–316.

Chou, R.Y., 1988. Volatility persistence and stock valuations: some empirical evidence using GARCH. J. Appl. Econ. 3, 279–294.

Davies, R.B., 1977. Hypothesis testing when a nuisance parameter is present only under the alternative. Biometrika 64, 247–254.

Davies, R.B., 1987. Hypothesis testing when a nuisance parameter is present only under the alternative. Biometrika 74, 3–43.

Diebold, F.X., 1986. Modelling the persistence of conditional variance: a comment. Econ. Rev. 5, 51–56.

Engle, R.F., 1982. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. Econometrica 50, 987–1007.

Engle, R.F., Bollerslev, T., 1986. Modelling the persistence of conditional variance. Econ. Rev. 5, 1–50.

Fama, E.F., French, K.R., 1988a. Business cycles and the behavior of metals prices. J. Fin. 43, 1075–1093.

Fama, E.F., French, K.R., 1988b. Commodity futures prices: some evidence on forecast power, premiums and the theory of storage. J. Bus. 60, 55–73.

Fong, W.M., 1997. Volatility persistence and switching ARCH in Japanese stock returns. Finan. Eng. Jpn. Mark. 4, 37–57.

French, K.R., Schwert, G.W., Stambaugh, R.F., 1987. Expected stock returns and volatility. J. Finan. Econ. 19, 3–30.

Gagnon, L., Lypny, G., 1997. The benefits of dynamically hedging the Toronto 35 stock index. Can. J. Admin. Sci. 14, 69–78.

Garcia, R., 1997. Asymptotic null distribution of the likelihood ratio test in Markov switching models. Int. Econ. Rev. 39, 763–788.

Gray, S.F., 1995. Managing Energy Price Risk. Risk Publication, London.

Gray, S.F. (1996a), An analysis of conditional regime-switching models. Working Paper, Duke University.

Gray, S.F., 1996b. Modelling the conditional distribution of interest rates as a regime switching process. J. Finan. Econ. 42, 27–62.

Hamilton, J.D., 1989. A new approach to the economic analysis of non-stationary time series and the business cycle. Econometrica 57, 357–384.

Hamilton, J.D., Susmel, R., 1994. Autoregressive conditional heteroskedasticity and changes in regime. J. Econ. 64, 307–333.

Hansen, B., 1992. The likelihood ratio test under non-standard conditions: testing the Markov switching model of GNP. J. Appl. Econ. 7, 61–82.

Kroner, K.F., Kneafsy, K. and Claessens, S. (1993), Forecasting volatility in commodity markets, World Bank Policy Research Working Paper, No. 1226, Washington, D.C.

Lamoureux, C.G., Lastrapes, W.D., 1990. Persistence in variance, structural change, and the GARCH model. J. Bus. Econ. Stat. 8, 225–243.

Litzenberger, R.H., Rabinowitz, N., 1995. Backwardation in oil futures markets: theory and empirical evidence. J. Fin. 50, 517–1546.

Nelson, D.B., 1991. Conditional heteroskedasticity in asset returns: a new approach. Econometrica 59, 347–370.

Ng, V.K., Pirrong, S.C., 1994. Fundamentals and volatility: storage, spreads, and the dynamics of metal prices. J. Bus. 67, 203–230.

Samuelson, P., 1965. Proof that properly anticipated prices fluctuate randomly. Indus. Manage. Rev. 6, 41–49.

Serletis, A., 1992. Maturity effects in energy futures. Energy Econ. 4, 150–157.

Verleger, P.K., 1994. Adjusting to Volatile Energy Prices. Institute for International Economics, Washington.

Wall Street Journal (1995), 1994 year-end review of markets and finance, 3 January, p. 34.

Webb, R.I., 1987. A note on volatility and pricing of futures options during choppy markets. J. Futures Mark. 7, 333–337.

Williams, J.C., Wright, B.D., 1991. Storage and Commodity Markets. Cambridge University Press, Cambridge.