Energy Economics 31 (2009) 779–788

Contents lists available at ScienceDirect

Energy Economics

journal homepage: www.elsevier.com/locate/eneco

Regime-switching stochastic volatility: Evidence from the crude oil market

Minh T. Vo[1]

College of Management, Metropolitan State University, 1501 Hennepin Avenue, Minneapolis, MN 55403, USA

A R T I C L E I N F O

Article history:
Received 22 August 2008
Received in revised form 30 April 2009
Accepted 1 May 2009
Available online 14 May 2009

JEL classification:
G10
G11
G32
C22

Keywords:
Crude oil price
Regime switching
Stochastic volatility
MCMC method
Forecasting

A B S T R A C T

This paper incorporates regime-switching into the stochastic volatility (SV) framework in an attempt to explain the behavior of crude oil prices in order to forecast their volatility. More specifically, it models the volatility of oil return as a stochastic volatility process whose mean is subject to shifts in regime. The shift is governed by a two-state first-order Markov process. The Bayesian Markov Chain Monte Carlo method is used to estimate the models. The main findings are: first, there is clear evidence of regime-switching in the oil market. Ignoring it will lead to a false impression that the volatility is highly persistent and therefore highly predictable. Second, incorporating regime-switching into the SV framework significantly enhances the forecasting power of the SV model. Third, the regime-switching stochastic volatility model does a good job in capturing major events affecting the oil market.

1. Introduction

Oil is one of the most important production factors in an economy. Over the period of 1998–2002, the U.S. spent about 7% of its GDP on energy, half of that on oil (Regnier, 2007). It is well-known that the crude oil markets sometimes appear quite calm and other times highly volatile. Modeling oil volatility is very important for at least three reasons. First, large movements in oil prices often have great impacts on the economy. History shows that most U.S. post World War II recessions were preceded by sharp increases in crude oil prices. The conventional explanation for this phenomenon is that a higher oil price will lower GDP growth by increasing production costs. Rotemberg and Woodford (1999) estimate that a 10% increase in oil price leads to an average GDP decline of 2.5% five or six quarters later. In addition, Bernanke (1983) and Pindyck (1991) show that large oil price movements increase uncertainty about future prices and thus cause delays in business investments. Given its significant impact on the U.S. economic activity, the Federal Reserve explicitly uses the volatility of oil as an input in formulating its monetary policy.

Second, volatility, in general, is an important variable for pricing derivatives, whose trading volume has quadrupled in recent years. Furthermore, volatility is needed to construct optimal hedge ratios to hedge against risk. Third, in making efficient econometric inference about the mean of a variable, we need a correct specification of its volatility. Therefore, understanding oil price behavior in order to forecast its volatility is a very important task for both monetary policymakers and practitioners in financial markets. It has attracted the attention of both academics and practitioners for the last three decades or so.

In oil economics, empirical studies, such as Abosedra and Laopodis (1997), Morana (2001), Bina and Vo (2007), suggest that crude oil price time series, like other financial time series, are featured by fat tail distribution, volatility clustering, asymmetry and mean reversion. Some studies also find strong evidence of regime switching. For instance, Wilson et al. (1996) find that there are structural changes in volatility in daily returns of oil futures due to exogenous factors such as severe weather conditions, political turmoil, and changes in OPEC oil policies. Fong and See (2002) find strong evidence of regime shifts in volatility of the West Texas Intermediate daily oil futures prices.

The autoregressive conditional heteroskedasticity (ARCH) models provide a breakthrough in modeling time-varying conditional variance. For comprehensive survey of the ARCH-type models, the reader is referred to Bollerslev et al. (1992), Bollerslev et al. (1994), Bauwens et al. (2006a), and Bauwens et al. (2006b). This type of

E-mail address: minh.vo@metrostate.edu.

[1] I would like to thank Ken Zapp, Allen Bellas, Narasimha Paravastu, Tim Delmont, Cyrus Bina, Liang Ding, Velayoudom Marimoutou, the referee, and seminar participants at Metropolitan State University Faculty Seminar for valuable comments which have helped me significantly improve the paper. I thank MaryAnn Johnson for clerical assistance, WTRG Economics for the permission to use the graph of oil prices in the paper. I would like to thank Dean Gary Seiler for support and encouragement. All remaining errors are my own responsibility.

0140-9883/$ – see front matter
doi:10.1016/j.eneco.2009.05.001

model assumes that the conditional variance is a deterministic linear function of past squared innovations and past conditional variances. Using several different univariate and multivariate models of ARCH-type, Sadorsky (2006) finds that the single-equation GARCH outperforms more sophisticated models in forecasting volatility of petroleum futures. Fong and See (2002) show that Markov switching GARCH models do a good job in capturing oil-related events, and outperform GARCH models in out-of-sample forecast.

Taylor (1986), however, contends that the volatility process should be driven by economic forces, not the past movements of prices as assumed in ARCH-type models. The stochastic volatility (SV) models are, therefore, developed under this belief. Surveys of the SV framework can be found in Ghysels et al. (1996) and Shephard (1996). Unlike the ARCH-type counterpart, the SV models involve two innovation processes, one for the observable data and one for its unobservable conditional variance. The SV framework has proved to be very useful in option pricing and has drawn much attention in the financial econometric literature. For instance, Melino and Turnbull (1990) find that SV models provide more accurate European option prices than the Black-Scholes model.

Which model, ARCH-type or SV, is a better tool to model volatility of financial time series? There is no definitive answer to this question. However, several studies find evidence in favor of SV. For instance, Kim et al. (1998) show that the basic SV model gives better in-sample fit than ARCH-type models. Geweke (2005) finds that both perform well in periods of low and sustained volatility; however, the SV model provides superior forecasts in periods of volatility jumps. Jacquier et al. (1994) find that compared to GARCH, the SV model yields a better and more robust description of the autocorrelation pattern of squared stock returns. Yu (2002) finds that the SV model outperforms the ARCH-type model in out-of-sample forecast. Given the evidence in favor of the SV framework and the fact that crude oil markets are highly volatile and exhibit volatility jumps, this paper chooses the SV framework to model the behavior of crude oil prices.

Several studies in literature, for example, Fong and See (2002), Morana (2001), among others, find that it is sufficient to model daily and weekly oil returns as a constant plus a random noise or just a random noise since the constant is not significantly different from zero. To ensure that the serial correlation is fully captured, a lag term is added into the model. More specifically, the oil return is modeled as an AR(1) process with stochastic variance. To take into account the regime change in volatility, the log variance is modeled as an AR(1) stochastic process whose unconditional mean is subject to shifts in regimes. There are two regimes under investigation: high volatility and low volatility regimes. They are governed by a first-order Markov process. To evaluate the Markov switching stochastic volatility (MSSV) model, its performance, in terms of goodness-of-fit and forecasting power, is compared to the constant variance AR(1) model, the GARCH model, the basic SV model, and the Markov switching (MS) model.

Estimating SV-type models is a challenge. These models do not have closed-form likelihood functions due to their latent structure of the variance. Thus, the maximum likelihood method cannot be used directly. Estimating the MSSV model is even harder because it involves two hidden levels, state and volatility, as opposed to just volatility in the SV model, or just state in the MS model. Several estimation methods have been proposed in the literature. The generalized method of moments (GMM) by Melino and Turnbull (1990), and Sorensen (2000) avoids the integration problem associated with the evaluation of the likelihood function. However, it is inefficient. The quasi-maximum likelihood (QML) of Harvey et al. (1994) is based on the maximization of the approximate likelihood function. It is simple to implement but inefficient if volatility proxies are non-Gaussian.

An alternative approach to GMM and QML is simulation-based methods. The indirect inference of Gourieroux et al. (1993) simulates the model using an auxiliary model and a calibration procedure. The efficient method of moments by Gallant et al. (1997) uses the score of the auxiliary model to improve the indirect inference method. The simulated maximum likelihood by Danielsson (1994), Durbin and Koopman (1997), and Sandmann and Koopman (1998) approximates the likelihood function by Monte Carlo simulation and then maximizes the approximate function.

In Bayesian framework, the Markov Chain Monte Carlo (MCMC) method uses the data augmentation principle to obtain a simulated sample from the posterior distributions of parameters and hidden states, given the available information. The theoretical underpinning of this method is the Hammersley-Clifford theorem which characterizes a joint distribution by its complete conditional distributions. Thus, MCMC method iteratively samples from these conditional distributions. When these distributions can be sampled directly, the algorithm is also known as Gibbs sampling. Jacquier et al. (1994) show that in SV framework, the MCMC method is superior to both QML and GMM.

Although MCMC is considered as one of the best estimation methods for SV-type models, it is computation-intensive and hard to implement because of the multi-dimensional integration problems in posterior distribution calculations. Nevertheless, this issue has been overcome by the availability of more powerful computers. Given its superiority, I choose the MCMC method to estimate the models.

The empirical results indicate that for all volatility models under investigation, thick tail is no longer present after we standardize the residuals by the estimated latent volatility. Among these models, the MS model fits the data the best. Heteroskedasticity is completely removed. Jarque-Bera and Kolmogorov-Smirnov tests for normality do not reject the null hypothesis of normality of the standardized residuals. I also find that volatility persistence in the SV model significantly reduces when regime shift is taken into account. In particular, the estimated persistence parameter in the SV model is 0.9578 while it is 0.5218 in the MSSV model. This result implies that the half life of a volatility shock, the time it takes for a shock to decay half of its initial value, is 16 weeks in the SV model while it is 1 week in the MSSV model. Thus, the MSSV model yields more reasonable result.

The paper also shows that the MSSV model does a good job in capturing major events in the market. For instance, it shows that the crude oil market switches to high volatility regime when Saudi Arabia abandoned its role as a swing producer in 1986, during the 1990 Gulf War, in the unusual cold spell in winter 1996, during the Asian financial crisis in 1997, after the terrorist attack on September 11, 2001, during the U.S. invasion of Iraq in 2003, etc.

In terms of forecasting power, the MSSV model provides more accurate out-of-sample forecast than the constant variance, GARCH (1,1), SV and MS models. This demonstrates that incorporating regime-switching into the SV framework will improve the forecasting power. In in-sample forecast, the result is mixed and dependent on the evaluation measure.

The remainder of the paper is organized as follows. Section 2 provides some background to the MS and SV framework. Section 3 describes the data set under investigation. Section 4 presents the empirical results. Section 5 compares models. Section 6 concludes the paper.

2. Model background and literature

This section discusses some basic models in the SV and MS framework which have been used extensively in literature. The first one is the SV model which is the most popular model in modeling time-varying volatility besides ARCH-type models. It is also the building block for other, more complex, models in the SV family. I then move on to present the MS model which is used to account for structural breaks in the time series. Finally, I discuss the MSSV model, which results from incorporating the MS model into the SV framework.

2.1. Stochastic volatility model

The SV model is parsimonious and yet flexible. It has proved useful in modeling a wide range of time series, including short-term interest rates, exchange rates, stock markets, etc. In the SV model, volatility is subject to a source of innovation which may or may not be correlated with those that drive returns. This is in direct contrast to ARCH-type models, which specify variance as a deterministic, often linear, function of lagged squared forecast errors and lagged conditional variance. The correlation between the return and the volatility processes can be used to capture volatility asymmetry. In this paper, I assume that there is no correlation between the two.[2]

In continuous-time, the SV model assumes that the time series follows a geometric Brownian motion and its log-variance follows an Ornstein-Uhlenback process. In discrete time, the model assumes the time series, having been demeaned, is a martingale while its log-variance follows an AR(1) process. More specifically, let r_t be the (demeaned) financial time series to be modeled, under SV framework, we have:

$$r_t = exp\left(\frac{h_t}{2}\right)u_t, u_t \sim IID(0,1),$$
$$h_t - \mu = \phi(h_{t-1} - \mu) + \sigma\varepsilon_t, \varepsilon_t \sim IID(0,1),$$

where h_t is the log-variance of r_t, μ is the unconditional mean of h_t, ϕ is the autoregressive term of the volatility process, and σ is the standard deviation of h_t. The standardized return and volatility innovations, u_t and ε_t, may or may not be correlated. Their correlation can be used to create volatility asymmetry (Hull and White, 1987). This modeling immediately leads to fat tail distribution for return, a stylized fact of most financial time series. The autoregressive term ϕ in the volatility process introduces persistence. It has to be less than one for the volatility process to be stationary.

A common finding in both ARCH-type and SV models is high persistence in the conditional variance. For instance, Chou (1988), French et al. (1987), Baillie and DeGennaro (1990), Fong (1997) report the persistent parameter above 0.9 for weekly stock returns. Brenner et al. (1996) and Anderson and Lund (1997) estimate that to be 0.82 and 0.98, respectively, for the weekly 3-month US T-Bill yields. High persistence implies that shocks to the conditional variance do not die out quickly and therefore, current information has significant impact on future volatility.

However, Diebold (1986), Lamoureux and Lastrapes (1990), among others, argue that in the presence of structural breaks, the persistent parameter could be overestimated and the MS model may be more suitable to model volatility. We now turn to the MS framework.

2.2. Markov switching model

The MS model, introduced by Hamilton (1988) and Hamilton (1989), assumes that the distribution of a variable is known conditioned on a particular regime occurring. At any point in time, the economy can be in any of a finite number of regimes. However, since regime is unobservable, the econometrician must draw statistical inference regarding the likelihood that each regime occurs at any point in time. The regime variable is assumed to follow the first-order Markov process with the transition probability from regime i to regime j defined as $p_{ij} = P(S_t = j | S_{t-1} = i)$. Thus, with n regimes, we need to estimate $n(n-1)$ transition probability parameters. It is possible to allow transition probabilities to evolve over time.

The MS model has been quite effective in modeling a wide range of economic time series, such as, stock market, business cycle, exchange rates, short-term rates, etc. Hamilton and Susmel (1994) fit Markov switching ARCH model to estimate stock volatility and find that the model provides a better fit to the data and better forecasts. Fong and See (2002) show that regime shifts are present in the crude oil futures price data and dominate GARCH effects. Aug and Bakaert (2002) find that MS models forecast interest rates better than single-regime models. Smith (2002) finds that either MS or SV model, but not both, is sufficient to fit the interest rate data; however, the MS model provides better interest rate forecasts. Kalimipalli and Susmel (2004) find that in short-term interest rate forecast, in-sample results strongly favor the regime-switching SV model as opposed to the single-state SV and GARCH models. However, out-of-sample results are mixed.

2.3. Markov-switching stochastic volatility model

The MSSV model incorporates MS model into the SV framework in an attempt to capture structural changes in the volatility process which cause the overestimation of the persistent coefficient. This model, thus, allows us to better understand the time series behavior when there is a jump in volatility due to economic forces and/or extreme market conditions.

Assuming the return follows an AR(1) process with stochastic variance. The MSSV model can be described as follows:

$$\begin{aligned} & r_t - m = \beta(r_{t-1} - m) + exp\left(\frac{1}{2}h_t\right)u_t, \quad u_t \sim IID(0,1) \\ & h_t - \mu_{s_t} = \phi\left(h_{t-1} - \mu_{s_{t-1}}\right) + \sigma\varepsilon_t, \quad \varepsilon_t \sim IID(0,1) \\ & \mu_{s_t} = \mu + \gamma s_t, \quad \gamma > 0\, s_t = \{0,1\} \\ & prob(s_t = j | s_{t-1} = i) = p_{ij}, \end{aligned} \tag{1}$$

where r_t is the return, m is the unconditional mean of return, h_t is the natural log of the conditional variance of the return, and ϕ measures the degree of persistence of volatility. The parameter μ_{s_t}, which is a function of the latent state S_t, is the stationary mean of h_t. μ_{s_t} follows a two-state ergodic discrete first-order Markov process as described in Hamilton (1988). The latent state can assume two possible values, 0 and 1, with the higher state corresponding to the higher stationary mean (i.e., $\mu_0 < \mu_1$). The sensitivity of μ_{s_t} to the latent state is represented by γ which is constrained to be positive. For identification reason, each regime must correspond to at least one data point. Shocks to the mean and volatility, μ_t and ε_t, are assumed to be independent. The parameter σ is the standard deviation of volatility shocks. The transition probability from state i to state j is represented by p_{ij}. Thus, the two-state transition probability matrix 2×2 governs the dynamics of the transition from one state to the other.

Model (1) nests both the SV model and the MS model. When γ is equal to zero, it reduces to the SV model. When both ϕ and σ are equal to zero, it becomes the MS model.

Table 1
Descriptive statistics.

	Price (\$)	r_t(%)	$\varepsilon_t = r_t - a - br_{t-1}$
Mean	28.08	6.06	0
Standard deviation	16.68	224.28	223.43
Skewness	1.84	−0.18	−0.17
Kurtosis	5.77	6.09	6.07
Jacque-Bera (p-value)		463.562 (0.0000)	456.5232 (0.0000)
Kolmogorov-Smirnov (p-value)		0.5226 (0.0000)	0.5162 (0.0000)
Ljung-Box (1) (p-value)		8.72 (0.0031)	13.12 (0.0003)
Ljung-Box (12) (p-value)		46.48 (0.0000)	298.01 (0.0000)
Ljung-Box (24) (p-value)		69.00 (0.0000)	418.58 (0.0000)
LB-ARCH (24) (p-value)			418.94 (0.0000)

[2] The SV model with correlation between the return and volatility has been estimated. It turns out that the correlation estimate is not significantly different from zero. For brevity, the result is not reported.

Table 2
Parameter estimates of the SV model.

Parameter	$\hat{R}$	Mean	STD	95% confidence interval
m	1.0006	1.5298	2.8194	(−3.9685, 7.0140)
β	1.0003	0.1406	0.0304	(0.0803, 0.2000)
μ	1.0012	10.6628	0.2207	(10.3250, 11.1800)
ϕ	1.0027	0.9578	0.0145	(0.9270, 0.9843)
σ	1.0007	0.2158	0.0293	(0.1624, 0.2802)

The table presents the parameter estimates along with their potential scale reduction factor scale R of the following SV model:

$$r_t - m = \beta(r_{t-1} - m) + exp\left(\frac{1}{2}h_t\right)u_t, \;\; u_t \sim N(0,1),$$
$$h_t - \mu = \phi(h_{t-1} - \mu) + \sigma\varepsilon_t, \;\; \varepsilon_t \sim N(0,1),$$
$$cov(u_t, \varepsilon_t) = 0.$$

Estimating the MSSV model involves the estimation of the mean parameters $\{m, \beta\}$, the volatility parameters $\{\phi, \sigma, \mu, \gamma\}$, the transition probability parameters $\{p_{11}, p_{22}\}$ (note that $p_{12} = 1 - p_{11}$ and $p_{21} = 1 - p_{22}$), and two latent variables h_t and s_t. Jacquier et al. (1994) show that maximum likelihood based methods often fail when the likelihood function is complex. The quasi-maximum likelihood is inefficient. Therefore, in this paper, I use the Bayesian MCMC method for estimation. This method is appropriate in models with latent components. The reader is referred to Chib (2001) for a survey on MCMC.

Let $\Omega = [m\,\beta\,\phi\,\sigma\,\mu\,\gamma]'$ be the vector of parameters, $H_t = [h_1\, h_2 \ldots h_t]'$ and $S_t = [s_1\, s_2 \ldots s_t]'$ be the vectors of latent variables, and $R_t = [r_1\, r_2 \ldots r_t]'$ be the vector of observed data. A full Bayesian model comprises the joint prior distribution of Ω, S_t, H_t, and R_t. Bayesian inference is then based on the joint posterior distribution of unobservables (Ω, H_t, S_t) given the data R_t. Let $f(.)$ be the probability density function. Using Bayes' theorem, we have:

$$f(H_t, S_t, \Omega \,|\, R_t) \propto f(R_t \,|\, H_t, \Omega) f(H_t \,|\, S_t, \Omega) f(S_t \,|\, \Omega) f(\Omega).$$

The Gibbs sampling algorithm is used to draw the conditional distributions $f(S_t|H_t, \Omega, R_t)$ $f(H_t|S_t, \Omega, R_t)$ and $f(\Omega|H_t, S_t, R_t)$ to generate samples from the joint posterior distribution. Based on the samples, I construct a 95% confidence interval and the standard error for each estimated parameter. I then use the samples generated to estimate the density function for each parameter.

3. Data

The data set is obtained from the historical crude oil price database of the US Department of Energy. It consists of a weekly spot price time series of the West Texas Intermediate (WTI) crude oil contracts traded on the New York Mercantile Exchange. The contract is denominated in 1,000 US barrels (42,000 gallons) of light sweet crude oil. The sample spans from January 3, 1986 to January 4, 2008 for a total of 1149 observations. Return is calculated as the difference in consecutive natural log prices. The return is then annualized for the analysis.

Table 1 presents descriptive statistics of the data set. The annualized return is quite volatile. Its sample mean is 6.06% while the sample standard deviation is 224.28%. The small but negative skewness of −0.18 indicates that the return distribution is slightly skew to the left. The kurtosis of 6.09 (this value is 3 for normal distribution) implies that return distribution has thicker tails than normal. Furthermore, both Jacque-Bera and Kolmogorov-Smirnov tests strongly reject the null hypothesis of normality of the return r_t and the residual ε_t in the regression of r_t on a constant and its lag.

Ljung-Box portmanteau tests at various lags suggest that both r_t and ε_t are highly autocorrelated. The last row is the Ljung-Box statistic for the squared residual ε_t^2 at various lags. The null hypothesis of no ARCH effects is strongly rejected by the data.

In short, the preliminary analysis shows that the return time series exhibits thick tails, autocorrelation, and heteroskedasticity. Thus, the SV framework is appropriate for modeling.

4. Empirical results

To examine the power of the SV and the MS framework in modeling the behavior of the crude oil price, I fit the SV, MS and MSSV models using the data set described above and compare their performance in terms of goodness-of-fit and forecasting power.

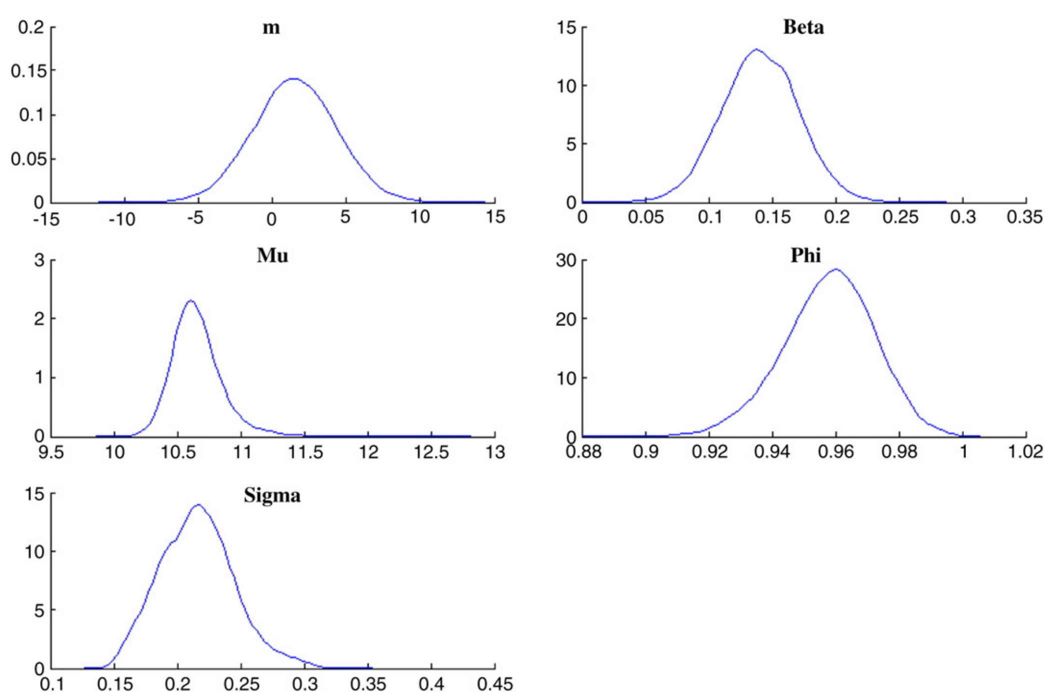

Fig. 1. Posterior probability densities for parameters in the SV model.

4.1. Stochastic volatility model

Setting γ equals to zero in Model (1), we get the following SV model:

$$\begin{aligned} & r_t - m = \beta(r_{t-1} - m) + exp\left(\frac{1}{2}h_t\right)u_t, u_t \sim IID(0,1), \\ & h_t - \mu = \phi(h_{t-1} - \mu) + \sigma\varepsilon_t, \varepsilon_t \sim IID(0,1), \\ & cov(u_t, \varepsilon_t) = 0. \end{aligned} \tag{2}$$

In this model, the return r_t follows an AR(1) process with stochastic volatility. The standardized mean and variance innovations, μ_t and ε_t, are assumed to be normal. Instead of estimating m and β in the mean equation by ordinary least square and then using the residuals from the regression to fit the variance equation as in Pagan and Schwert (1990), Ball and Torous (1999), Kalimipalli and Susmel (2004), I estimate both equations jointly using the Bayesian MCMC method. The prior distributions of m and μ are assumed to be normal. That of $\frac{1}{\sigma^2}$ is gamma. In order for r_t and h_t to be stationary, β and ϕ are constrained to the interval $(-1,1)$ by setting $\beta = 2\beta^* - 1$ and $\phi = 2\phi^* - 1$, where $\beta^*$and ϕ^* have beta prior. These are natural conjugate priors, so the posterior densities will have the same form as those of the priors. This will facilitate the sampling process. Since there is no prior knowledge about these parameters, I use very large variances for these prior distributions.

To investigate the sensitivity of the posterior results to the initial values, I generate multiple chains with starting points sampled from the over-dispersed prior distributions. I then check whether they converge to the same distribution. More specifically, I randomly generate three sets of initial values from the prior distributions, from which three Markov chains are generated. I then use the Gelman and Rubin's (1992) approach to test if these chains converge to the same posterior distribution.

Table 3
Parameter estimates of the MS model.

Parameter	$\hat{R}$	Mean	STD	95% confidence interval
m	1.0000	2.2484	2.8454	(−3.2825, 7.7970)
β	1.0002	0.1473	0.0291	(0.0897, 0.2039)
σ_0	1.0061	173.7340	4.8238	(164.4, 183.2)
σ_1	1.0130	434.2444	32.6184	(375.1, 502.85)
p_{11}	1.0002	0.9867	0.0052	(0.9744, 0.9948)
p_{22}	1.0041	0.9140	0.0311	(0.8431, 0.9634)

The table presents the parameter estimates along with their potential scale reduction factor scale $\hat{R}$ of the following MS model:

$$\begin{aligned} & r_t - m = \beta(r_{t-1} - m) + \sigma_{s_t}u_t,\ \ u_t \sim N(0,1) \\ & \sigma_{s_t} = \mu + \gamma s_t,\ \ \gamma > 0 \text{ and } s_t = \{0,1\} \\ & prob(s_t = j | s_{t-1} = i) = p_{ij}. \end{aligned}$$

The Gelman-Rubin test uses the potential scale reduction factor $\hat{R}$ to assess the convergence of the chain:

$$\hat{R} = \frac{c+1}{c}\left(\frac{n-1}{n} + \frac{B/n}{W}\right) - \frac{n-1}{cn},$$

where c is the number of chain, n is the number of iteration after the burn-in period, B/n is the between-sequence variance, and W is the within-sequence variance. Gelman and Rubin (1992) show that if $\hat{R}$ is close to one then the multiple chains converge to the same distribution and the estimate is not sensitive to the initial value.

Table 2 reports the posterior results. $\hat{R}$ is very close to one for every parameter, indicating that all three chains converge to a stationary posterior distribution and so the results are not sensitive to the initial values. Furthermore, we observe that all parameters but m are significantly different from zero. The persistence parameter φ is very close to unity (0.9578), suggesting that a shock to volatility is highly persistent. Half-life of a volatility shock, the time it takes for the volatility shock to decay half of its initial value, calculated by $-\ln(2)/\ln(\varphi)$, is about 16 weeks. Fig. 1 presents the posterior densities for the parameters.

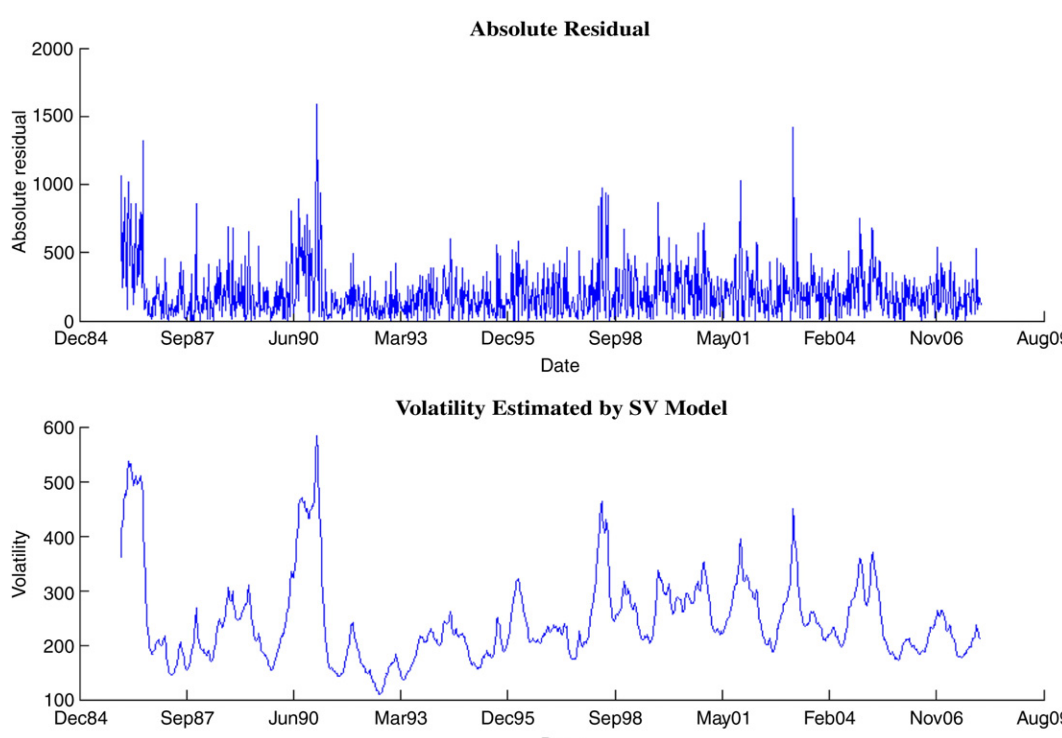

Fig. 2. Volatility estimated from the SV model versus the absolute value of the residual in the mean equation.

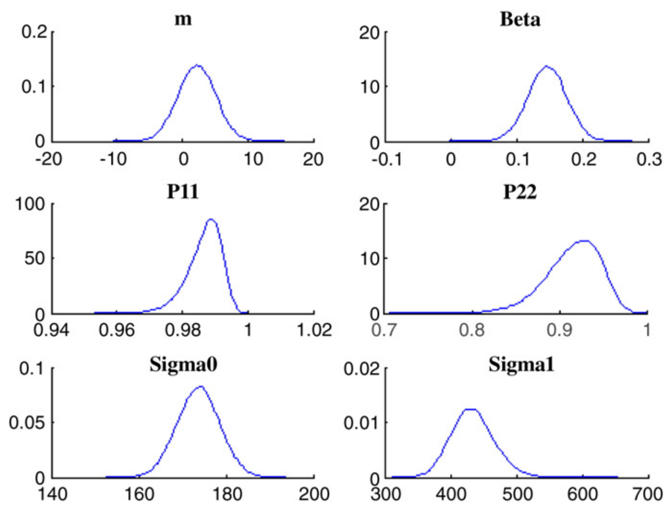

Fig. 3. Posterior probability densities for parameters in the MS model.

Fig. 2 plots the volatility time series estimated from the SV model and the absolute value of the residual in the mean equation. With the naked eyes, we see that the estimated volatility fits quite well the pattern of the residual.

4.2. Markov switching model

Setting ϕ and σ equal to zero in Model (1) and re-parameterize, we get the following MS model:

$$\begin{aligned} &r_t - m = \beta(r_{t-1} - m) + \sigma_{s_t} u_t, \; u_t \sim N(0,1) \\ &\sigma_{s_t} = \mu + \gamma s_t, \; \gamma > 0 \text{ and } s_t = \{0,1\} \\ &prob(s_t = j | s_{t-1} = i) = p_{ij}. \end{aligned} \tag{3}$$

Like the SV model in the previous section, the return r_t still follows an AR(1) process but its volatility switches between the low value σ_0 and the high value σ_1 depending on the state s_t.

Table 4
Parameter estimates of the MSSV model.

Parameter	$\hat{R}$	Mean	Standard deviation	95% confidence interval
m	0.9999	2.4005	2.8646	(−3.101, 8.0155)
β	1.0036	0.1285	0.0294	(0.071, 0.1862)
μ_0	1.0004	10.4378	0.0683	(10.3, 10.57)
μ_1	1.2761	17.1809	1.9277	(13.96, 21.12)
ϕ	1.1767	0.5218	0.0670	(0.3842, 0.6436)
p_{11}	1.0431	0.9977	0.0022	(0.9918, 0.9998)
p_{22}	1.0043	0.5955	0.1859	(0.2039, 0.9058)
σ	1.0718	0.6631	0.0565	(0.5588, 0.7797)

The table presents the parameter estimates along with their potential scale reduction factor scale R of the following MSSV model:

$$\begin{aligned} &r_t - m = \beta(r_{t-1} - m) + exp\left(\frac{1}{2}h_t\right)u_t, \; u_t \sim N(0,1) \\ &h_t - \mu_{s_t} = \phi\left(h_{t-1} - \mu_{s_{t-1}}\right) + \sigma\varepsilon_t, \; \varepsilon_t \sim N(0,1) \\ &\mu_{s_t} = \mu + \gamma s_t, \; \gamma > 0\, s_t = \{0,1\} \\ &prob(s_t = j | s_{t-1} = i) = p_{ij}, \\ &cov(u_t, \varepsilon_t) = 0. \end{aligned}$$

Model (3) is estimated using the Bayesian MCMC method with Gibbs sampling algorithm. The prior distribution of m is assumed to be normal, while that of p_{11} and p_{22} is beta ($p_{12} = 1 - p_{11}$, and $p_{21} = 1 - p_{22}$) to ensure that they are constrained in the interval [0,1]. The autoregressive parameter β has to be in the interval $(-1,1)$ in order for AR(1) model to be stationary. For that reason, I set $\beta = 2\beta^* - 1$, where β^* has prior beta distribution.

To take into account the fact that σ_1 is greater than $\sigma_0 (\gamma > 0)$, I use the following hierarchical structure for their priors:

$$\begin{aligned} &\sigma_0 \sim Uniform(lower, m_0), \\ &\sigma_1 \sim Uniform(m_1, upper), \\ &lower \sim Uniform(a, b), \\ &m_0 \sim Uniform(lower, b), \\ &m_1 \sim Uniform(m_0, b), \\ &upper \sim Uniform(m_1, b), \end{aligned} \tag{4}$$

where the parameters a and b are chosen such that they do not influence the result. I choose a and b to be 0 and 1000 respectively. Table 3 reports the estimation results.

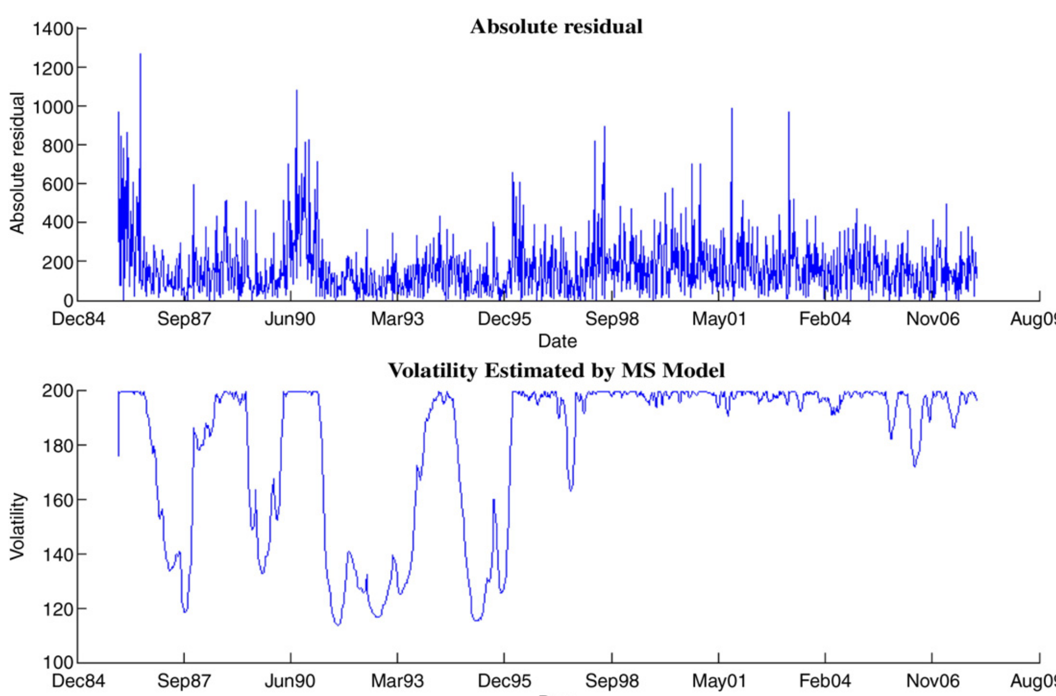

Fig. 4. Volatility estimated from the MS model versus the absolute value of the residual in the mean equation.

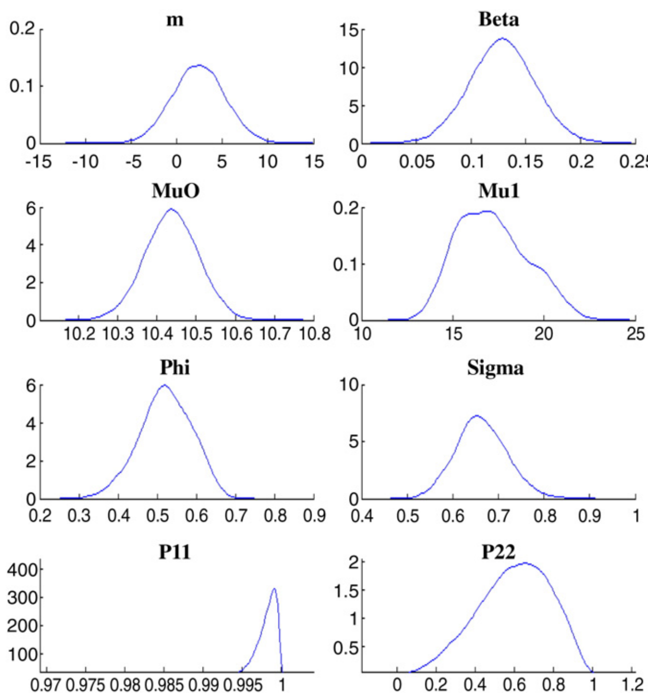

Fig. 5. Posterior probability densities for parameters in the MSSV model.

As in the SV models, all parameter estimates except m are highly significant. The autoregressive term estimate β is very close to that in the case of SV model (0.1473 vs. 0.1406). The volatility in the high regime is more than twice that in the low regime. Fig. 3 plots the posterior densities of the parameters estimated.

Fig. 4 plots the volatility estimated from the MS model and the absolute residual in the mean equation.

4.3. Markov switching stochastic volatility model

This section presents the estimates of the following MSSV model:

$$r_t - m = \beta(r_{t-1} - m) + exp\left(\frac{1}{2}h_t\right)u_t,\ \ u_t \sim N(0,1)$$
$$h_t - \mu_{s_t} = \phi\left(h_{t-1} - \mu_{s_{t-1}}\right) + \sigma\varepsilon_t,\ \ \varepsilon_t \sim N(0,1)$$
$$\mu_{s_t} = \mu + \gamma s_t,\ \ \gamma > 0\, s_t = \{0,1\}$$
$$prob(s_t = j | s_{t-1} = i) = p_{ij},$$
$$cov(u_t, \varepsilon_t) = 0.$$

The model is estimated by the Bayesian MCMC method. To express the initial vague knowledge about the parameters, I use non-informative priors as in previous sections. Table 4 reports the results.

We observe that all parameter estimates except for m are significantly different from zero as in the previous cases. However, the persistence parameter Φ is significantly reduced to 0.5218 from 0.9578 in the SV model. This estimate is comparable to 0.51 reported in Fong and See (2003) in the regime-switching GARCH framework. Thus, half-life of a volatility shock is about one period (one week), meaning that it takes a week for the volatility shock to decay half of its initial value. Recall that this measure is 16 weeks in SV model. This result confirms earlier findings in the literature that changes in regime might create an artificially high

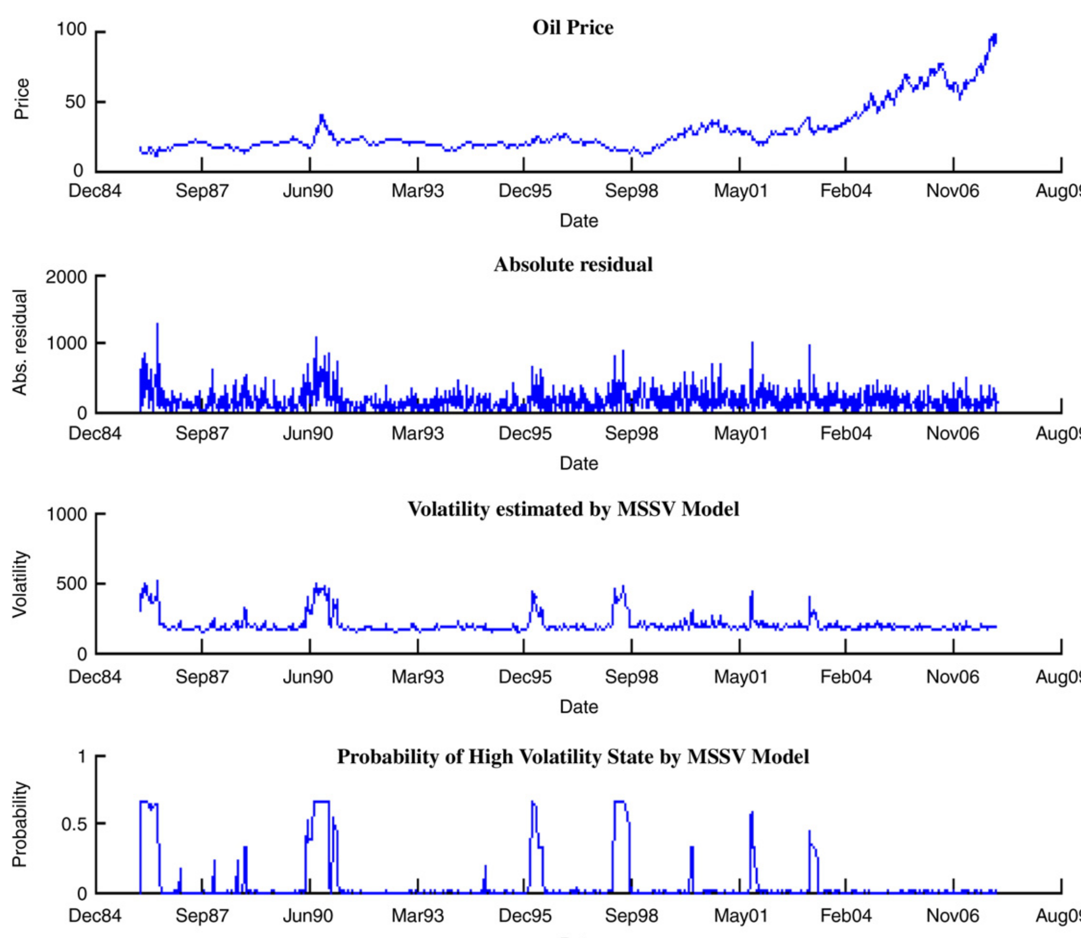

Fig. 6. Prices, returns, corresponding latent volatilities, and states.

persistence in volatility. Fig. 5 plots the posterior density of each parameter.

Fig. 6 plots the oil prices, the absolute residual, the latent volatility and the probability of the high volatility state indicated by the MSSV model. It shows that the price has had a consistent up-trend since 2002. The absolute residual time series displays volatility clustering, a common feature of general financial time series. There are a few instances where volatility surges to high level, corresponding to the high state of volatility. The first instance is around 1986 when Saudi Arabia, the dominant member of OPEC stopped acting as a swing producer. Indeed, as a swing producer, Saudi Arabia cut output when supply in the market was high and raised output when it was low in order to keep crude oil price in an "optimal" range. However, some factors, such as the deregulation of the U.S. oil market and the considerable increase in the output from non-OPEC oil producers have caused Saudi Arabia to lose revenue and market share significantly. This forced Saudi Arabia to abandon this strategy in 1986 and as a result oil prices collapsed.

The second instance is during the gulf war from August 1990 to February 1991. Others include the exceptionally cold winter spell in 1996 in the northern hemisphere, the Asian crisis in 1997, the terrorist attack on September 11, 2001, and the U.S. invasion of Iraq in 2003. These events either caused the oil price to rise or fall but raised the volatility in the market. Thus, the MSSV model does quite well in capturing most major events in the market. Fig. 7, extracted from WTRG Economics, features oil prices from 1970 to 2009 in 2008 dollars and major oil-related events. The reader is referred to the WTRG website, http://www.wtrg.com/prices.htm, for the complete historical analysis of oil price.

5. Comparing models

5.1. Model checking

This section assesses the goodness-of-fit of each model. To that end, I test whether the volatility indicated by the model is sufficient to explain the time series behavior. More specifically, I test whether the residual in the return equation, having been standardized by the corresponding volatility in the variance equation, is normal, using the Jarque-Bera and Kolmogorov-Smirnov tests. Table 5 presents the test results.

We see that the residuals of all models have means close to zero and standard deviations close to one. In addition, the skewness of the distributions in all models is close to zero, indicating that they are almost symmetric. Furthermore, the kurtosis measures in all three models are close to 3, indicating that the distributions of standardized residuals do not have thick tails. For the MS model, the null hypothesis of normality cannot be rejected at 5% level of significance (p-values are 0.3564 in Jacque-Bera test and 0.0771 in Kolmogorov-Smirnov test). However, the tests reject the null hypothesis in the SV and MSSV models with p-values of 0.0149 and 0.0000 respectively in Jacque-Bera test, and 0.0371 and 0.0262 respectively in Kolmogorov-Smirnov test. Even so, their deviation from normality is less pronounced (The Jarque-Bera and Kolmogorov-Smirnov statistics are smaller than 456.52 and 0.5226 respectively when volatility is not modeled). Thus, this result indicates that the MS model would fit the data the best. To further check the models, in the next section, I will compare their forecasting power, both in-sample and out-of-sample.

5.2. Forecasting performance

The first step in evaluating the forecasting performance of a model is to obtain realized volatility. There are a number of ways in the literature to obtain realized volatility (see, for example, Merton, 1980; Perry, 1982; Akgiray, 1989; Ding et al., 1993). This paper uses the approach of Merton (1980) and Perry (1982) by simply summing squared daily returns in a week to get the realized volatility of that week. More specifically,

$$\sigma_t^2 = \sum_{i=1}^{n_t} r_i^2, \qquad (5)$$

where r_i is the daily return on day i and n_t is the number of trading days in week t. To evaluate the forecasting performance, I use three

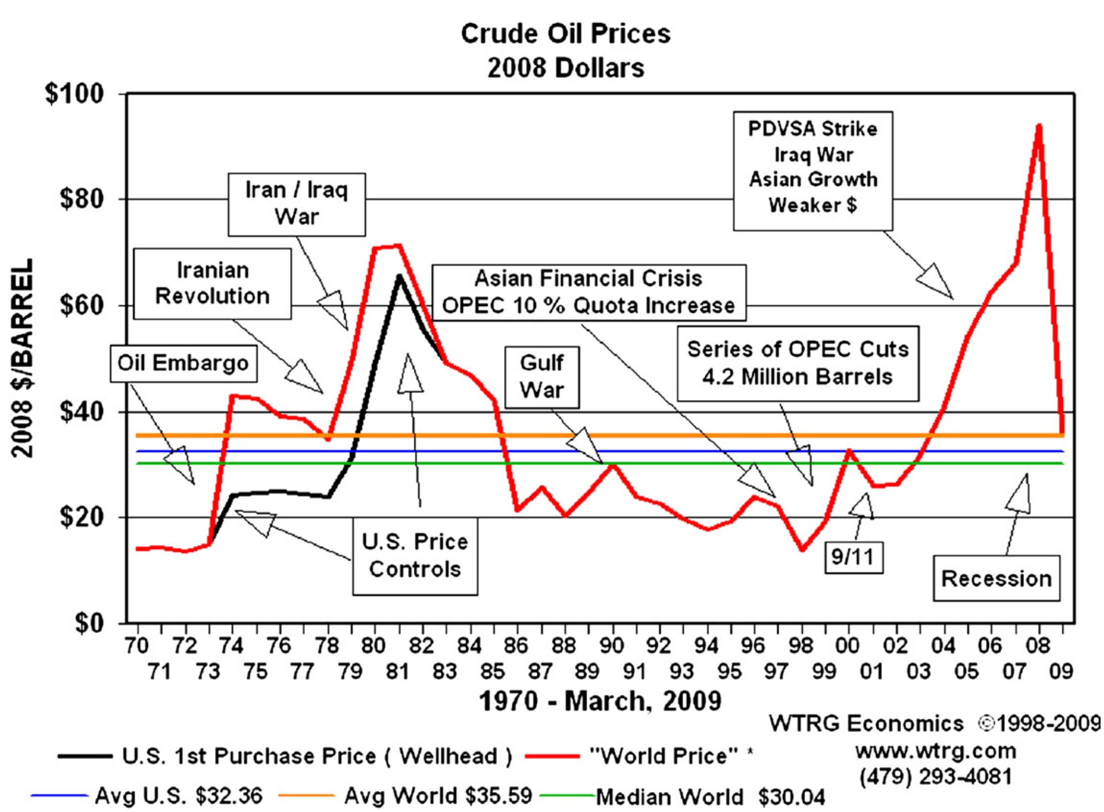

Fig. 7. Crude oil prices 1970–2009 (source: WTRG Economics).

most popular metrics: the root mean square error (RMSE), the mean absolute error (MAE), and the Theil-U statistic. They are given by:

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\hat{\sigma}_i^2 - \sigma_i^2\right)^2}, \tag{6}$$

$$\mathrm{MAE} = \frac{1}{N}\sum_{i=1}^{N}\left|\hat{\sigma}_i^2 - \sigma_i^2\right|, \tag{7}$$

$$\mathrm{Theil{-}U} = \frac{\sum_{i=1}^{N}\left(\hat{\sigma}_i^2 - \sigma_i^2\right)^2}{\sum_{i=1}^{N}\left(\hat{\sigma}_{\mathrm{base},i}^2 - \sigma_{\mathrm{base},i}^2\right)^2}, \tag{8}$$

where $\hat{\sigma}_i^2$ is the forecast volatility in week i indicated by the model, σ_i^2 is the realized volatility in week i, calculated by Eq. (5). N is the number of weeks in the sample. The subscript "base" in the Theil-U metric indicates the base model which is the AR(1) model with constant variance in this case. Thus, for the constant variance AR(1), the Theil-U statistic is always equal to unity. In all cases, the lower the metrics, the better is the model in terms of forecasting power.

To conduct the forecasting experiment, I follow three steps. First, the models are estimated for each sample. Second, based on the estimated models, the volatilities over the next 10 weeks are forecast. Third, the metrics are calculated for comparison.

In addition to the full sample, 01/03/86–01/04/08, I consider two subsamples. The first one is from 01/03/1986 to 01/08/1999, and the second one is from 01/08/1999 to 01/04/2008. The first subsample includes two major events, Saudi Arabia no longer a swing producer, and the Gulf War in 1990. The second subsample includes the unusual cold winter 1996 in the northern hemisphere, and the Asian financial crisis in 1997, the terrorist attack on September 11, 2001, and the U.S. invasion of Iraq. For more extensive evaluation, I also include the constant variance and the GARCH(1,1) models in the comparison. Table 6 presents the results.

For the in-sample forecast, under the MAE measure, the MS model is the best for the whole sample. This is consistent with the test result in the previous section where the null hypothesis of normality of standardized residual cannot be rejected only in the MS model. The MSSV model is rank second. The GARCH(1,1) and the constant variance models are ranked third and fourth. The SV model is the weakest. However, for subsample 1, the MSSV model is the best and the MS is at the bottom. For the second subsample, the best is the MS model, followed by the MSSV.

Under the other two measures, RMSE and Theil-U, the constant variance AR(1) model performs the best in the whole sample and the first subsample. However, in the second subsample, the MSSV is the best. Thus the in-sample forecast result is mixed and dependent on the evaluation metric. We cannot tell which model is the best.

In the out-of-sample forecast, the MSSV model consistently produces the best forecasts under all three evaluation metrics. The MS model comes second in the full sample and the first subsample. The SV model is second in the second subsample.

Table 5
Diagnostic test on standardized residuals in the return equations of SV, MS and MSSV models.

	SV	MS	MSSV
Mean	0.0294	0.0334	0.0292
Standard deviation	0.9566	0.9973	0.9094
Skewness	−0.1870	−0.0847	−0.1390
Kurtosis	2.7853	2.8857	2.2378
Jarque-Bera (p-value)	8.8901 (0.0149)	1.9979 (0.3564)	31.4563 (0.0000)
Kolmogorov-Smirnov (p-value)	0.0415 (0.0371)	0.0375 (0.0771)	0.0433 (0.0262)

Table 6
Forecasting performance of competing models.

Sample	Model	MAE		RMSE		Theil-U	
		Value	Rank	Value	Rank	Value	Rank
In-sample							
Full sample	Const. variance	0.0931	4	0.0933	1	1	1
01/03/86 to	GARCH(1,1)	0.0912	3	0.1105	2	1.4027	2
01/04/08	SV	0.1239	5	0.1541	4	2.7280	4
	MS	0.0823	1	0.1106	3	1.4052	3
	MSSV	0.0855	2	0.2570	5	7.5876	5
Subsample 1	Const. variance	0.0938	4	0.0940	1	1	1
01/03/86 to	GARCH(1,1)	0.0932	3	0.1261	4	1.7996	4
01/08/99	SV	0.0846	2	0.1205	3	1.6433	3
	MS	0.1079	5	0.1659	5	3.1148	5
	MSSV	0.0788	1	0.1168	2	1.5544	2
Subsample 2	Const. variance	0.0808	5	0.0811	3	1	3
01/08/99 to	GARCH(1,1)	0.0806	4	0.0810	2	0.9975	2
01/04/08	SV	0.0805	3	0.0811	3	1	3
	MS	0.0754	1	0.0852	5	1.1037	5
	MSSV	0.0768	2	0.0787	1	0.9417	1
Out-of-sample							
Full sample	Const. variance	0.0944	5	0.0944	5	1	5
01/03/86 to	GARCH(1,1)	0.0757	3	0.0758	3	0.6448	3
01/04/08	SV	0.0906	4	0.0907	4	0.9231	4
	MS	0.0747	2	0.0749	2	0.6295	2
	MSSV	0.0654	1	0.0654	1	0.4800	1
Subsample 1	Const. variance	0.0911	3	0.0914	3	1	3
01/03/86 to	GARCH(1,1)	0.1216	5	0.1222	5	1.7875	5
01/08/99	SV	0.1182	4	0.1184	4	1.6781	4
	MS	0.0818	2	0.0834	2	0.8326	2
	MSSV	0.0621	1	0.0635	1	0.4827	1
Subsample 2	Const. variance	0.0834	4	0.0834	4	1	4
01/08/99 to	GARCH(1,1)	0.08829	3	0.0829	3	0.9880	3
01/04/08	SV	0.0794	2	0.0795	2	0.9087	2
	MS	0.0936	5	0.0936	5	1.2596	5
	MSSV	0.0749	1	0.0749	1	0.08066	1

In short, even though the in-sample forecast result is mixed, the out-of-sample forecast is overwhelmingly in favor of the MSSV model. This result demonstrates that incorporating the MS model into the SV framework significantly improves the short-term forecasting power of the SV model.

6. Conclusions

This paper incorporates regime switching into the SV framework in an attempt to explain the behavior of crude oil price in order to forecast its volatility. More specifically, oil return is modeled as a stochastic volatility process whose mean is subject to shifts between two regimes. The regime is governed by a first-order Markov process. The model is then estimated by the Bayesian MCMC method with the Gibbs sampling algorithm. The main findings of the study are as follows. First, there is clear evidence of regime-switching in the oil market. Ignoring it will lead to a false impression that the volatility is highly persistent and therefore highly predictable. Second, the MSSV model does a good job in capturing most major events affecting the oil market. Third, incorporating regime-switching into the SV framework will significantly enhances the forecasting power of the SV model.

References

Abosedra, S.S., Laopodis, N.T., 1997. Stochastic behavior of crude oil prices: a GARCH investigation. Journal of Energy and Development 21, 283–291.
Akgiray, V., 1989. Conditional heteroscedasticity in time series of stock returns: evidence and forecast. Journal of Business 62, 55–80.
Anderson, T., Lund, J., 1997. Estimating continuous time stochastic volatility models of the short-term interest rates. Journal of Econometrics 77, 343–377.
Aug, A., Bakaert, G., 2002. Regime switches in interest rates. Journal of Business & Economic Statistics 20, 163–182.
Baillie, R.T., DeGennaro, R.P., 1990. Stocks returns and volatility. Journal of Financial and Quantitative Analysis 25, 203–215.

Ball, C., Torous, W.N., 1999. The stochastic volatility of short-term interest rates: some international evidence. Journal of Finance 56, 2339–2359.

Bauwens, L., Laurent, S., and Rombouts, J.V.K., 2006a, Regime switching GARCH models, CORE Discussion Paper.

Bauwens, L., Laurent, S., Rombouts, J.V.K., 2006b. Multivariate GARCH models: a survey. Journal of Applied Econometrics 21, 79–109.

Bernanke, B.S., 1983. Irreversibility, uncertainty, and cyclical investment. Quarterly Journal of Economics 98, 85–106.

Bina, C., Vo, M., 2007. OPEC in the epoch of globalization: an event study of global oil prices. Global Economy Journal 7.

Bollerslev, T., Chou, R.Y., Kroner, K.F., 1992. ARCH modeling in finance: a selective review of theory and empirical evidence. Journal of Econometrics 52, 5–59.

Bollerslev, T., Engle, R.F., Nelson, D.B., 1994. ARCH models. In: Engle, R., McFadden, D. (Eds.), Handbook of Econometrics. InNorth-Holland, New York.

Brenner, R.J., Harjes, R., Kroner, K., 1996. Another look at models of short-term interest rates. Journal of Financial and Quantitative Analysis 31, 85–107.

Chib, S., 2001. Markov chain Monte Carlo methods: computation and inference. In: Heckman, J.J., Leamer, E. (Eds.), Handbook of Econometrics. InNorth-Holland, Amsterdam.

Chou, R.Y., 1988. Volatility persistence and stock valuation: some empirical evidence using GARCH. Journal of Applied Econometrics 3, 279–294.

Danielsson, J., 1994. Stochastic volatility in asset prices: estimation with simulated maximum likelihood. Journal of Econometrics 64, 375–400.

Diebold, F., 1986. Comment on modeling the persistence of conditional variances. Econometric Review 5, 51–56.

Ding, Z., Granger, C.W.J., Engle, R.F., 1993. A long memory property of stock returns and a new model. Journal of Empirical Finance 1, 83–106.

Durbin, J., Koopman, S., 1997. Monte Carlo maximum likelihood estimation for non Gaussian state space models. Biometrika 84, 669–684.

Fong, W.M., 1997. Volatility persistence and switching ARCH in Japanese stock returns. Asia-Pacific Financial Market 4, 37–57.

Fong, W.M., See, K.H., 2002. A Markov switching model of the conditional volatility of crude oil futures prices. Energy Economics 24, 71–95.

Fong, W.M., See, K.H., 2003. Basic variations and regime shifts in the oil futures market. European Journal of Finance 9, 499–513.

French, K.R., Schwert, G.W., Stambaugh, R.F., 1987. Expected stock returns and volatility. Journal of Financial Economics 19, 3–30.

Gallant, A.R., Hsieh, D., Tauchen, G., 1997. Estimation of stochastic volatility models with diagnostics. Journal of Econometrics 81, 159–192.

Gelman, A., Rubin, D.B., 1992. Inference from iterative simulation using multiple sequences. Statistical Science 7 (4), 457–511.

Geweke, J., 2005. Contemporary Bayesian econometrics and statistics. John Wiley & Sons, Hoboken, NJ.

Ghysels, E., Harvey, A.C., Renault, E., 1996. Stochastic volatility. In: Rao, C.R., Maddala, G.S. (Eds.), Statistical Models in Finance. InNorth-Holland, Amsterdam.

Gourieroux, C., Monfort, A., Renault, E., 1993. Indirect inference. Journal of Applied Econometrics 8, 85–118.

Hamilton, J.D., 1988. Rational expectations econometric analysis of changes in regime: an investigation of the term structure of interest rates. Journal of Economic Dynamics and Control 12, 385–423.

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. Journal of Econometrics 64, 307–333.

Harvey, A.C., Ruiz, E., Shephard, N., 1994. Multivariate stochastic variance models. Review of Economic Studies 61, 247–264.

Hull, J., White, A., 1987. The pricing of options on assets with stochastic volatilities. Journal of Finance 42, 281–300.

Jacquier, E., Polson, N.G., Rossi, P.E., 1994. Bayesian analysis of stochastic volatility models. Journal of Business and Economic Statistics 12, 371–389.

Kalimipalli, M., Susmel, R., 2004. Regime-switching stochastic volatility and short-term interest rates. Journal of Empirical Finance 11, 309–329.

Kim, S., Shephard, N., Chib, S., 1998. Stochastic volatility: likelihood inference and comparison with ARCH models. Review of Economic Studies 65, 363–393.

Lamoureux, C.G., Lastrapes, W.D., 1990. Persistence in variance, structural change, and the GARCH model. Journal of Business and Economic Statistics 8, 225–234.

Melino, A., Turnbull, S.M., 1990. Pricing foreign currency options with stochastic volatility. Journal of Econometrics 45, 239–265.

Merton, R., 1980. On estimating the expected return on the market: an explanatory investigation. Journal of Financial Economics 8, 323–361.

Morana, C., 2001. A semi-parametric approach to short-term oil price forecasting. Energy Economics 23, 325–338.

Pagan, A.R., Schwert, G.W., 1990. Alternative models for conditional stock volatility. Journal of Econometrics 45, 267–290.

Perry, P., 1982. The time-variance relationship of security returns: implications for the return-generating stochastic process. Journal of Finance 37, 857–870.

Pindyck, R., 1991. Irreversibility, uncertainty, and investment. Journal of Economic Literature 29, 110–148.

Regnier, E., 2007. Oil and energy price volatility. Energy Economics 29, 405–427.

Rotemberg, J.J., Woodford, M., 1999. Imperfect competition and the effects of energy price increases on economic activity. Journal of Money, Credit and Banking 28, 550–577.

Sadorsky, P., 2006. Modeling and forecasting petroleum futures volatility. Energy Economics 28, 467–488.

Sandmann, G., Koopman, S.J., 1998. Estimation of stochastic volatility models via Monte Carlo maximum likelihood. Journal of Econometrics 87, 271–301.

Shephard, N., 1996. Statistical aspects of ARCH and stochastic volatility. In: Cox, D.R., Hinkley, D.V. (Eds.), Time Series Models in Econometrics, Finance and Other Fields. InChapman & Hall, London.

Smith, D., 2002. Markov-switching and stochastic volatility diffusion models of short-term interest rates. Journal of Business & Economic Statistics 20, 183–197.

Sorensen, M., 2000. Prediction-based estimating functions. Econometrics Journal 3, 123–147.

Taylor, S.J., 1986. Modelling Financial Time Series. Wiley, New York.

Wilson, B., Aggarwal, R., Inclan, C., 1996. Detecting volatility changes across the oil sector. Journal of Futures Markets 16, 313–320.

Yu, J., 2002. Forecasting volatility in the New Zealand stock market. Applied Financial Economics 12, 193–202.