Yu-Min Lian (Taiwan), Szu-Lang Liao (Taiwan)

The volatility structure of oil futures market returns: an empirical investigation

Abstract

In this study, it is investigated the impact of suddenly structural breaks on estimated GARCH-type models with normal and heavy-tailed distributions for daily oil futures market returns. More specifically, the multiple structural breaks in return variance over the whole sample period are detected by the Inclán-Tiao's algorithm. The estimated results of the ICSS AR-GARCH models show that the volatility persistence decreases dramatically after controlling for such discrete breakpoints. The changing oil futures risk can be best captured by the ICSS AR-EGARCH-GED model. Specifically, the comparison of the in-sample model evaluation champions the AR-EGARCH-*t* model over competing models within each identified sub-period.

Keywords: oil futures market return, structural break, heavy-tailed distribution, in-sample model evaluation.
JEL Classification: C22, C51, G14.

Introduction

A variety of oil-linked derivatives, such as oil futures contracts, have been designed for hedging the oscillating risk in the oil market. This kind of speculative investment often follows a path of relative steady, disconnected by periods of greater market disturbance. This provides a problem for those trying to model the volatility dynamics. Previous studies investigate the modeling of changing volatility in various financial time series, especially stock market returns, foreign exchange rates, and so forth. Nevertheless, relatively little attention has been given to model oil futures returns in the context of volatility models. Due to the presence of non-normality in asset returns, which means that the asymmetric GARCH (e.g. EGARCH or TGARCH) models with heavy-tailed distributions may provide a better fitness to the data, as opposed to the symmetric GARCH models. Regarding the distributional properties of oil futures returns, we model the conditional variance using GARCH-type models with normal distribution, *t*-distribution, and the generalized error distribution (GED). Furthermore, currently more and more empirical evidences have revealed that the existence of structural breaks in financial time series can have serious implications on pricing-related derivatives. Without incorporating structural breaks into the analytical model may cause an overestimate of the volatility persistence in variance (Diebold, 1986; Lastrapes, 1989; Lamoureux and Lastrapes, 1990; Ewing and Malik, 2005). To detect the structural breaks in return variance, the iterated cumulative sums of squares (ICSS) algorithm proposed by Inclán and Tiao (1994) is considered to identify the presence of such breakpoints. This technique focuses on finding a statistically significant change in variance due to a breakpoint in the process that generates the volatility of the time series. After determining the aforementioned structural breaks by the ICSS algorithm, the resulting shifts can be incorporated into conditional variance of the analytical model in form of dummy variables for volatility analysis (Wilson et al., 1996; Aggarwal et al., 1999; Bracker and Smith, 1999; Malik, 2003; Malik and Hassan, 2004; Mansur et al., 2007; McMillan and Wohar, 2011; Huang, 2014).

With the increasing number of surprising events, the subprime crisis of 2008, for instance, has created large fluctuations in the oil market. Therefore, it is crucial to detect the volatility shifts adequately and model the future dynamics corresponding to the changing oil price according to actual market developments. After controlling for multiple structural breaks, we further add the resulting dummies to the GARCH models with three types of distributional specifications on the standardized residuals. In addition, we compare the in-sample model evaluation of the GARCH-type models in the full period and each sub-period identified using the Inclán and Tiao (1994) test. Empirical results are significant for the risk management of market participants.

Given the relatively few literature for modeling volatility changes in the oil futures market returns. In this study, we study the modeling of time-varying volatilities in oil futures returns under GARCH-type models by incorporating both structural breaks and heavy-tailed distributions generated by the oil futures market, which extends the classical ARCH/GARCH models for the oil futures price modeling. Accurately modeling volatility changes in oil futures market returns have significant implications for risk management and for determining dynamic hedging strategies, which is particularly important during unstable oil markets.

Yu-Min Lian, Corresponding author, Ph.D. in Banking and Financial Markets, Department of Money and Banking, College of Commerce, National Chengchi University, Taiwan.
Szu-Lang Liao, Professor, Department of Money and Banking, College of Commerce, National Chengchi University, Taiwan.

The remainder of this study is organized as follows. The next section illustrates the dynamic models. Section 2 presents the empirical results. The final section shows the conclusions of this study.

1. Methodology

1.1. The ICSS algorithm and identification of D_k. The iterated cumulative sums of squares (ICSS) algorithm, developed by Inclán and Tiao (1994), is used to detect discrete changes in variance of a time series. Let $C_k = \sum_{t=1}^{k} r_t^2$ be the cumulative sum of squares of a series of uncorrelated random variables, r_t, with mean 0 and variances σ_t^2, k = 1, …, T. Define the mean centered cumulative sum of squares as follows:

$$D_k = \frac{C_k}{C_T} - \frac{k}{T},\ k = 1,...,T\ ,$$

$$\text{with } D_0 = D_T = 0\,. \qquad (1)$$

For a series with homogeneous variance over the sample period, the D_k statistics oscillate around zero. On the contrary, when there is a sudden change in variance, the D_k value will exhibit a positive or negative drift away from zero. Inclán and Tiao (1994) calculate the critical values under the null hypothesis of constant variance from the asymptotic distribution of D_k. When the maximum of $|D_k|$ exceeds the critical value, the null hypothesis of no changes in variance is rejected. Denote the value of k at which $\max_k |D_k|$ is attained as k^*. If the maximum of $\sqrt{T/2}|D_k|$ is larger than the critical value of ±1.358 at the 5% level, then k^* is considered as an estimate of the breakpoint. The factor $\sqrt{T/2}$ is required to standardize the distribution. If a series has multiple breakpoints, the usefulness of the D_k function becomes doubtful because of the masking effect. Inclán and Tiao (1994) propose an iterative algorithm based on repeated applications of D_k on different segments of the series, dividing consecutively after a breakpoint is identified.

1.2. The GARCH models and distributional assumptions. The GARCH-type models are widely used in various branches of econometrics, especially in financial time series analysis. The simplest GARCH(1,1) model can be set as follows:

$$Y_t = X_t^{'}\theta + \varepsilon_t\,, \qquad (2)$$

$$\sigma_t^2 = w + \alpha\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2\,, \qquad (3)$$

where the mean equation given in (2) is written as a function of exogenous variables with an error term. The restrictions $w > 0$, α, and $\beta \geq 0$ in conditional variance equation (3) are imposed to insure a positive variance. An additional restriction is that both ARCH and GARCH models assume symmetry in the distribution of asset returns. It is well-known that many financial time series have non-normal distribution. Engle and Ng (1993) examine how negative shocks increase conditional volatility in stock market returns. These stock market returns are, like oil futures market returns, negatively skewed with heavy-tailed distributions. This suggests that asymmetric GARCH models might also be of value in capturing oil futures price movements. The so-called EGARCH (exponential GARCH) model was proposed by Nelson (1991). The specification for the conditional variance of the EGARCH(1,1) model is set to be:

$$\ln(\sigma_t^2) = w + \alpha\left|\frac{\varepsilon_{t-1}}{\sigma_{t-1}} - \sqrt{\frac{2}{\pi}}\right| + \beta\ln(\sigma_{t-1}^2) + \gamma\frac{\varepsilon_{t-1}}{\sigma_{t-1}}\,, \qquad (4)$$

where ε_t follows a generalized error distribution in equation (4). Note that the left-hand side is the log of the conditional variance. This implies that the leverage effect is exponential, rather than quadratic, and that forecasts of the conditional variance are guaranteed to be non-negative. The presence of leverage effects can be tested by the hypothesis that $\gamma < 0$. The impact is asymmetric if $\gamma \neq 0$. Alternative specification that is designed to capture the increasing volatility from asymmetric shocks is the TGARCH model. The TGARCH (threshold GARCH) or GJR-GARCH model were introduced independently by Zakoïan (1994) and Glosten, Jaganathan, and Runkle (1993). The specification for the conditional variance of the TGARCH(1,1) model is given by:

$$\sigma_t^2 = w + \alpha\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2 + \gamma\varepsilon_{t-1}^2 I_{t-1}\,, \qquad (5)$$

where I_{t-1} = 1 if ε_{t-1} < 0 and 0 otherwise. In this model, good news, ε_{t-1} > 0, and bad news. ε_{t-1} < 0, have differential effects on the conditional variance; good news has an impact of α, while bad news has an impact of $\alpha + \gamma$. If $\gamma > 0$, bad news increases volatility, and we say that there is a leverage effect. If $\gamma \neq 0$, the news impact is asymmetric.

In order to capture the tail distributional characteristics of financial time series, it is essential to make distributional assumptions about the error term ε_t. There are three assumptions commonly employed when working with GARCH models: normal distribution, *t*-distribution, and the GED. Given a distributional setting, the GARCH models are typically estimated by the method of maximum likelihood. For example, for the GARCH(1,1) model with conditionally normal errors, the log-likelihood function of sample is given by the following:

$$\ln L(\psi) = -\frac{T}{2}\ln(2\pi) - \frac{1}{2}\sum_{t=1}^{T}\ln\sigma_t^2 - -\frac{1}{2}\sum_{t=1}^{T}(Y_t - X_t'\theta)^2 / \sigma_t^2, \quad (6)$$

where σ_t^2 is the conditional variance of the error term ε_t. Under t-distribution, the log-likelihood function of this type of sample is of the following form:

$$\ln L(\psi) = -\frac{T}{2}\ln\left\{\frac{\pi(k-2)\Gamma(k/2)^2}{\Gamma\left[(k+1)/2\right]^2}\right\} - -\frac{1}{2}\sum_{t=1}^{T}\ln\sigma_t^2 - \frac{(k+1)}{2}\sum_{t=1}^{T}\ln\left[1 + \frac{(Y_t - X_t'\theta)^2}{\sigma_t^2(k-2)}\right], \quad (7)$$

where the degree of freedom $k > 2$ controls the tail behavior. The t-distribution approaches the normal as $k \to \infty$. Under GED, the log-likelihood function of this type of sample can be written as follows:

$$\ln L(\psi) = -\frac{T}{2}\ln\left[\frac{\Gamma(1/v)^3}{\Gamma(3/v)(v/2)^2}\right] - -\frac{1}{2}\sum_{t=1}^{T}\ln\sigma_t^2 - \sum_{t=1}^{T}\ln\left[\frac{\Gamma(3/v)(Y_t - X_t'\theta)^2}{\sigma_t^2\Gamma(1/v)}\right]^{v/2}, \quad (8)$$

where the tail parameter $v > 0$. The GED is a normal distribution if $v = 2$, and fat tail if $v < 2$.

1.3. Empirical model setting. In this section, we use dummy variables representing volatility changes identified by the ICSS algorithm into the GARCH-type processes. The specification for the ICSS AR(p)-GARCH(1,1) model is therefore set to be the following:

$$r_t = \varphi r_{t-p} + \varepsilon_t, \quad (9)$$

$$\sigma_t^2 = w + \alpha\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2 + \sum_{l=1}^{n} d_l D_l, \quad (10)$$

After controlling for detected breakpoints, the specification for the EGARCH(1,1) and TGARCH(1,1) processes are respectively given by:

$$\ln(\sigma_t^2) = w + \alpha\left|\frac{\varepsilon_{t-1}}{\sigma_{t-1}} - \sqrt{\frac{2}{\pi}}\right| + \beta\ln(\sigma_{t-1}^2) + +\gamma\frac{\varepsilon_{t-1}}{\sigma_{t-1}} + \sum_{l=1}^{n} d_l D_l, \quad (11)$$

and

$$\sigma_t^2 = w + \alpha\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2 + \gamma\varepsilon_{t-1}^2 I_{t-1} + \sum_{l=1}^{n} d_l D_l, \quad (12)$$

where D_1, ..., D_n in conditional variance equations (10)-(12), are the set of dummy variables taking a value of one from each breakpoint of variance onwards and zero otherwise. All models are estimated by the method of maximum likelihood under three types of distributional assumptions that the errors with normal, t, and GEDs.

2. Empirical results

2.1. Data description. The daily data for the Light-Sweet oil futures contracts are from Datastream and cover the period from 1 August 1997 to 31 July 2007 (2608 observations). First differences in natural logarithms of the price levels are employed in all models. The top panel of Figure 1 shows the daily oil futures data in level form, while the bottom panel of Figure 1 shows the return series and a high degree of variability in returns. From the bottom panel of Figure 1, the series exhibits a large number of volatility, while showing a tendency towards a constant mean. Of course, it is necessary that the data be mean reverting. Otherwise, the variance tends to infinity as the number of observations approaches infinity, presenting the t-values undependable and inducing spurious results.

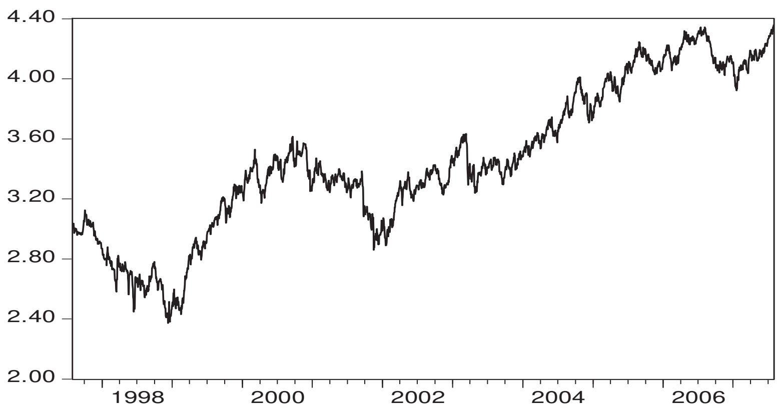
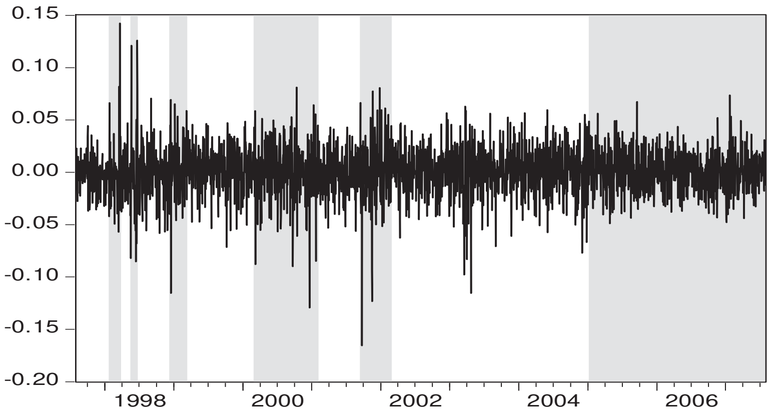


Fig. 1a. Time series (top panel) and logarithmic returns (bottom panel) of daily oil futures prices



Notes: The daily price data are based on oil futures contracts traded on the New York Mercantile Exchange (NYMEX). The contracts expire four times per year (March, June, September, and December). Three month contracts are used to construct a continuous series. In order to avoid any expiration effects, the new contract start a week before the expiration of the former contract. The shaded areas in the bottom panel of Figure1 indicate periods of changing volatility detected using the ICSS algorithm.

Fig. 1b. Time series (top panel) and logarithmic returns (bottom panel) of daily oil futures prices

The dataset used in the descriptive analysis consists of the daily oil futures prices, and summary statistics of the return series are presented in Table 1. The skewness and kurtosis suggest a leptokurtic distribution with negatively skewed returns in the oil futures market. The Jarque-Bera statistics represent that the return series are not normally distributed. Therefore, both the skewness and the tail behavior of the data should be better captured by the asymmetric GARCH models with heavy-tailed distributions, which are designed to model asymmetry and fat tail in this study.

Table 1. Summary statistics of the return series

	Light-Sweet oil futures returns
Mean	5.20E-04
Variance	5.36E-04
Skewness	-0.283***
Kurtosis	6.722***
Jarque-Bera	1540.458***
Observations	2608

Notes: The futures data are from Datastream and cover the period from 1 August 1997 to 31 July 2007. Jarque-Bera is the test for normality. ***, **, and * represent statistical significance at the 0.01, 0.05, and 0.10 levels, respectively.

Table 2 reports the Augmented Dickey-Fuller and Phillps-Perron unit root test statistics for the logs of prices and daily oil futures return series. As reported in Table 2, ADF(C,T,0) is the Augmented Dickey-Fuller unit root test with constant, trend and lags of 0. ADF(1) and ADF(0) are the Augmented Dickey-Fuller unit root test with 1 and 0 lags, respectively. PP(C,T,14) is the Phillps-Perron test with constant, trend and lags of 14. PP(17) is the Phillps-Perron test with 17 lags. Results from Table 2 show that the ADF and PP unit root test statistics for the logs of prices are not able to reject the null hypothesis that the existence of a unit root at the 0.05 significance level. Furthermore, the ADF and PP unit root test statistics for the first-differenced daily return series are all well below the critical values at the 0.01 significance level, indicating a strong rejection of the presence of a unit root. Therefore, the daily oil futures returns are first difference stationary and proceed with the proposed tests.

Table 2. Unit root tests for the logs of daily oil futures prices and return series

	ADF(C,T,0)	ADF(C,T,0)	PP(C,T,14)
Logs of oil futures prices	-3.205*	-3.205*	-3.009
	ADF(1)	ADF(0)	PP(17)
Oil futures returns	-37.584***	-50.902***	-51.155***

Notes: The lags selections based on AIC and SIC are in the second column and the third column, respectively. The lags selection based on Newey-West Bandwidth using Bartlett Kernel in the last column. The 0.01, 0.05, and 0.01 critical values for ADF(C,T,0) and PP(C,T,14) are -3.961, -3.411, and -3.127, respectively. The 0.01, 0.05, and 0.01 critical values for ADF(1), ADF(0) and PP(17) are -2.565, -1.940, and -1.616, respectively. The null hypothesis for the ADF and PP tests is the presence of a unit root. ***, **, and * represent significance at the 0.01, 0.05, and 0.01 levels, respectively.

2.2. Report of detected breakpoints. The bottom panel of Figure 1 shows that the presence of time-varying volatility clustering phenomena and many spikes in the data, with more negative than positive outliers. This is consistent with the significant negatively skewed and excess kurtosis reported in Table 1. Furthermore, the non-normality of the return series takes the use of the ICSS algorithm to detect structural breaks in variance. There are eleven structural breaks in return series detected by the ICSS algorithm. We divide the full period into twelve sub-periods for the return series to provide evidence of unstable GARCH process and hence changing volatilities. Table 3 reports such breakpoint dates, along with selected news that are associated with volatility shifts in the return series. Obviously, there is a great deal of variance within each sub-period and the suddenly discrete volatility jumps at the breakpoints. The empirical evidence indicates that the return variances are not constant over the tested period. This is a confirmation that more exactness of the ICSS algorithm is imperative for modeling asset returns. Due to the residual GARCH effects may in a volatility analysis, the ICSS algorithm may not capture all of the variance effects. Hence, a more intact analysis would think about both kinds of impacts. Correspondingly, we examine the GARCH effects, as well as the existence of suddenly discrete volatility shifts.

Table 3. Structural breaks detected by the ICSS algorithm

Dates	Days	Wall Street news on oil futures
01/22/1998	125	Oil futures drop on bearish inventory data.
03/25/1998	44	Oil rockets 13% on OPEC cutback--03/24.
05/18/1998	38	Oil futures drop as glut continuous built.
06/23/1998	26	Oil futures pass $15 a barrel--06/24.
12/09/1998	121	Oil prices swing on report--12/08.
03/11/1999	66	Oil gains ahead of a March 23 OPEC meeting.
02/29/2000	253	Oil reaches high; other sectors languish-02/28.
02/02/2001	243	Oil futures soars to an eight-week high.
09/13/2001	159	Terrorist attack--09/11.
02/26/2002	118	Dow industrials surge 263 points--03/01.
01/06/2005	747	Oil futures jump more than $2--01/07.
07/31/2007	668	
Total	2608	

Notes: Dates are the ending days for the sub-period. The last column comes from the Wall Street Journal on ProQuest Newspapers.

2.3. Modeling the oil futures returns. We identify the best-fitting specification of conditional mean equation by Box-Jenkins procedures. The partial autocorrelation function suggests that the AR(2) process would be appropriate for the return series. Table 4 reports that the $Q(15)$ and $Q(20)$ statistics are not significant, indicating there is no serial correlation in returns. The $Q^2(15)$ and $Q^2(20)$ statistics are significant, indicating statistically significant serial correlation in squared returns, which motivates us to model the conditional heteroskedasticity.

Table 4. Serial correlation tests for AR(2) process of daily oil futures returns

Serial correlation and ARCH tests for the AR(2) process	
Q(15)	10.120
Q(20)	16.367
Q^2(15)	94.412***
Q^2(20)	102.390***

Notes: Q(n) and Q^2(n) are the Ljung-Box test statistics for the 15th and 20th order serial correlation in the ordinary and squared ordinary returns, respectively. ***, **, and * represent statistical significance at the 0.01, 0.05, and 0.10 levels, respectively.

As noted above, we model the oil futures return using the AR(2)-GARCH(1,1) process. One way of further examining the distribution of the residuals is to plot the quantiles. Figure 2 indicates that the QQ-plots of standardized residuals for the AR(2)-GARCH(1,1) model with normal distribution. If the residuals are normally distributed, the points in the QQ-plots should lie alongside a straight line. As shown in Figure 2, the QQ-plots show that it is primarily large negative shocks that are driving the departure from normality. Because of the possibility that the appearance of non-normality in residuals, we further examine the distributional characteristics by GARCH-based models under heavy-tailed distributions in the next subsection.

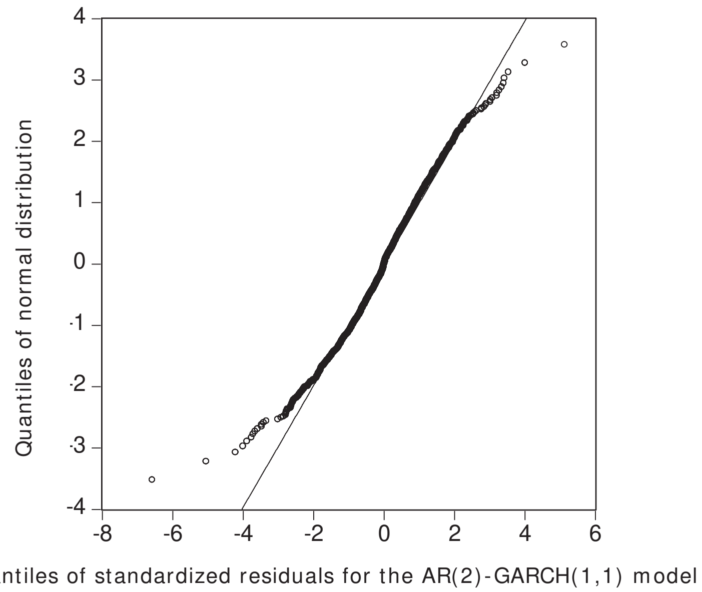


Fig. 2. QQ-plots of standardized residuals for the AR(2)-GARCH(1,1) model with normal distribution

2.4. Comparison of the estimated results. Time variation in the second moments for the full sample is modeled using the AR-GARCH models under three types of distributional assumptions, the estimated results of which are reported in Table 5. As shown in Table 5, the GARCH coefficient, β, in each model ranges from 0.902 to 0.979. These estimates are consistent with those found in models of stock returns. In addition, all the results exhibit significantly high levels of volatility persistence (close to an I-GARCH process) for the return series. The shock parameter, γ, in the AR-EGARCH/TGARCH models with three types of distributional assumptions are estimated to be negative and positive, respectively. Again, it is consistent with the existence of significant leverage effect in stock return models, indicating that shocks are greater than expected raise variance. The degrees of freedom parameters, κ, and v, in the AR-GARCH models with heavy-tailed distributions are estimated to be $4 < k < \infty$ and $0 < v < 2$, respectively. The estimates suggest that the heavy-tailed distributions of the standardized errors depart significantly from normality. The estimated results show that the significant non-normality in return series. According to the comparison of log-likelihood values in each model, the AR-EGARCH-GED model without volatility shifts fits best for the series. Finally, the test statistics, Q, and Q^2, are not significant at the 15th and 20th lags, so there is little evidence of serial correlation and remaining ARCH effects in standardized and squared standardized residuals, respectively. The empirical results in Table 5 suggest that the AR-EGARCH-GED specification is more appropriate than competing models for the modeling of oil futures returns and that the volatility analysis should incorporate time-varying second moments.

Table 5. Estimated results of the AR-GARCH models with normal, *t*, and GED

	AR(2)-GARCH(1,1)	AR(2)-EGARCH(1,1)	AR(2)-TGARCH(1,1)
ϕ	-0.045** (-2.119)	-0.045** (-2.198)	-0.044** (-2.110)
	-0.025 (-1.350)	-0.025 (-1.353)	-0.024 (-1.296)
	-0.007 (-0.435)	-0.008 (-0.467)	-0.007 (-0.409)
ω	2.10E-05*** (2.886)	-0.265*** (-3.053)	1.85E-05*** (2.929)
	8.69E-06** (2.547)	-0.201*** (-3.093)	1.02E-05*** (2.852)
	1.04E-05** (2.490)	-0.231*** (-3.082)	1.17E-05*** (2.751)
α	0.059*** (3.034)	0.083*** (3.484)	0.027 (1.406)
	0.023*** (3.871)	0.060*** (3.867)	0.007 (0.896)
	0.030*** (4.155)	0.070*** (3.943)	0.011 (1.150)
β	0.902*** (32.990)	0.973*** (95.834)	0.915*** (37.657)
	0.960*** (90.064)	0.979*** (125.609)	0.958*** (87.886)
	0.949*** (72.386)	0.976***(108.802)	0.949*** (72.279)
γ		-0.053*** (-3.087)	0.046** (2.264)
		-0.036*** (-3.567)	0.029** (2.492)
		-0.042*** (-3.893)	0.034** (2.520)
κ	5.875*** (8.882)	6.088*** (8.626)	6.089*** (8.723)
v	1.241*** (29.750)	1.252*** (30.066)	1.251*** (29.933)
log L	6189.877	6195.382	6195.338
	6260.627	6263.416	6263.672
	6264.532	6267.831	6267.665
Q(15)	6.227	7.061	6.613
	7.056	7.452	7.383
	8.487	9.131	8.888
Q(20)	9.875	11.247	10.329
	11.201	11.796	11.502
	12.514	13.488	12.923
Q^2(15)	13.686	16.411	14.007
	22.835*	20.487	19.120
	18.947	18.251	16.796
Q^2(20)	16.823	19.729	17.441
	25.863	23.537	22.179
	22.352	21.471	20.175

Notes: Within each cell the estimate with normal distribution is the top parameter, *t*-distribution is the middle parameter, and GED is the bottom parameter. *z*-statistics are in parentheses. κ and v are degrees of freedom parameters for the *t*-distribution and GED, respectively. log L represents the log-likelihood values. Q(n) and Q^2(n) are the Ljung-Box test for the 15th and 20th order serial correlation in standardized and squared standardized residuals, respectively. ***, **, and * represent statistical significance at the 0.01, 0.05, and 0.10 levels, respectively.

Recall the previous findings that there are eleven structural breaks detected by the ICSS algorithm in Section 2.2. Here, we employ dummy variables representing such breakpoints into the GARCH-type processes. As shown in Table 6, the ARCH coefficient, α, estimated from all models under three types of distributional assumptions are not statistically significant. The values of coefficient β become smaller after controlling for structural breaks. Both the test results suggest that the ARCH effects vanish and the degree of persistence is significantly reduced after controlling for such detected breakpoints. This means that, the volatility persistence overestimated by the GARCH-type models can be explained by structural breaks. Furthermore, all of the discrete breakpoints bring about sizable shifts in the intercept term, w, and that these shifts often lead to substantial changes in variance across regimes, that is, non-stationary of the variance. To consider whether the addition of volatility shift dummy variables leads to a statistically superior model specification relative to competing models, we compare the log-likelihood values of each model, discovering the superiority of the extended ICSS AR-GARCH models. Specifically, the ICSS AR-EGARCH-GED model fits best for the return series. Thus, a volatility analysis should be modeled to accommodate both GARCH effects and volatility changes.

Table 6. Estimated results of the ICSS AR-GARCH models with normal, *t*, and GED

	ICSS AR(2)-GARCH(1,1)	ICSS AR(2)-EGARCH(1,1)	ICSS AR(2)-TGARCH(1,1)
ϕ	-0.037* (-1.954)	-0.033* (-1.723)	-0.036* (-1.877)
	-0.025 (-1.337)	-0.021 (-1.143)	-0.025 (-1.302)
	-0.012 (-0.680)	-0.009 (-0.535)	-0.011 (-0.634)
ω	8.28E-05** (2.236)	-3.644** (-2.572)	9.46E-05** (2.270)
	9.80E-05* (1.885)	-3.428*** (-3.533)	8.28E-05** (2.508)
	8.85E-05* (1.819)	-3.586*** (-3.241)	8.54E-05** (2.282)
α	0.030 (1.545)	0.041 (0.805)	-0.005 (-0.210)
	0.018 (1.183)	0.011 (0.297)	-0.009 (-0.560)
	0.023 (1.288)	0.025 (0.602)	-0.008 (-0.403)
β	0.599*** (3.660)	0.566*** (3.346)	0.570*** (3.413)
	0.541** (2.362)	0.591*** (5.153)	0.616*** (4.357)
	0.567** (2.560)	0.574*** (4.399)	0.595*** (3.632)
γ		-0.096*** (-2.784)	0.067** (1.989)
		-0.108*** (-3.981)	0.066** (2.349)
		-0.104*** (-3.554)	0.066** (2.083)
κ	7.479*** (6.770)	7.497*** (6.681)	7.600*** (6.724)
v	1.343*** (26.530)	1.348*** (26.018)	1.348*** (26.502)
log *L*	6248.475	6254.583	6251.741
	6285.022	6291.154	6288.044
	6293.279	6298.516	6295.834
Q(15)	6.376	7.1939	6.9590
	6.911	7.7388	7.1678
	7.493	8.3153	7.8499
Q(20)	10.908	11.621	11.397
	11.492	12.117	11.529
	12.061	12.707	12.247
Q^2(15)	25.381**	26.123**	25.107**
	26.883**	23.423*	20.530
	25.251**	23.375*	21.557*
Q^2(20)	25.951	26.918	25.768
	27.682*	24.411	21.365
	26.008	24.326	22.374

Notes: Within each cell the estimate with normal distribution is the top parameter, *t*-distribution is the middle parameter, and GED is the bottom parameter. *z*-statistics are in parentheses. κ and v are degrees of freedom parameters for the *t*-distribution and GED, respectively. log *L* represents the log-likelihood values. Q(n) and Q^2(n) are the Ljung-Box test for the 15th and 20th order serial correlation in standardized and squared standardized residuals, respectively. ***, **, and * represent statistical significance at the 0.01, 0.05, and 0.10 levels, respectively.

2.5. In-sample model evaluation. Summary statistics and in-sample RMSEs within the full period and each sub-period of daily oil futures returns are shown in Table 7. As shown in Table 3, the twelve sub-periods within the full period are determined by eleven structural breaks in return series using the ICSS algorithm. Obviously, the daily returns of the full period are negatively skewed with heavy-tailed. It is clear that the significant non-normality in the full period data. In itself, one might expect that the AR-EGARCH/TGARCH models with heavy-tailed distributions might yield superior modeling results. Indeed, the AR-GARCH-normal model has the highest RMSE. The AR-GARCH/EGARCH/TGARCH-GED models perform the best as well as exactly the same lowest RMSE (0.019169) for the series. This suggests that with considerable observations there is trivial to distinguish between these models. Thus, this study attempts to confirm that this conclusion holds for periods of changing volatility. Furthermore, within each sub-period of normally distributed data of Table 7, the AR-EGARCH/TGARCH models with heavy-tailed distributions perform poorly. However, there are six sub-periods (2^{nd}, 6^{th}, 7^{th}, 8^{th}, 10^{th}, and 11^{th}) in which the return series appear to be non-normally distributed. The AR-EGARCH-*t* model ranks first over competing models in the 2^{nd}, 6^{th}, 7^{th}, and 11^{th} sub-periods. Once more data become available, it is necessary to engage a more extensive analysis, not only with respect to the in-sample analysis but also in terms of the out-of-sample forecast.

Table 7. Summary statistics and in-sample RMSEs within the full period and each sub-period

	Full period	1^{st} sub-period	2^{nd} sub-period	3^{rd} sub-period	4^{th} sub-period	5^{th} sub-period	6^{th} sub-period
Numbers	2608	125	44	38	26	121	66
Variance	5.36E-04	2.29E-04	1.17E-03	4.28E-04	2.73E-03	5.08E-04	9.17E-04
Skewness	-0.283***	-0.018	1.875***	0.188	0.669	-0.130	-0.703**
Kurtosis	6.722***	3.619	8.664***	2.568	3.549	3.621	5.405***
AR(2)-GARCH(1,1) RMSE	0.019196 (9)	0.008236 (7)	0.066530 (2)	0.029657 (6)	0.075979 (1)	0.021328 (8)	0.032517 (7)
	0.019178 (5)	0.008194 (5)	0.066564 (3)	0.029336 (4)	0.077873 (4)	0.021282 (6)	0.032036 (3)
	0.019169 (1)	0.008161 (2)	0.067544 (8)	0.029686 (7)	0.077596 (3)	0.021194 (4)	0.032886 (9)
AR(2)-EGARCH(1,1) RMSE	0.019195 (8)	0.008205 (6)	0.067378 (6)	0.028506 (1)	0.077307 (2)	0.021247 (5)	0.032127 (4)
	0.019178 (5)	0.008164 (4)	0.066420 (1)	0.028543 (2)	0.081560 (5)	0.021364 (9)	0.031297 (1)
	0.019169 (1)	0.008161 (2)	0.067378 (6)	0.028561 (3)	0.085641 (7)	0.021301 (7)	0.032879 (8)
AR(2)-TGARCH(1,1) RMSE	0.019194 (7)	0.008266 (8)	0.066868 (4)	0.029759 (8)	0.083805 (6)	0.021182 (1)	0.031854 (2)
	0.019177 (4)	0.008363 (9)	0.067369 (5)	0.029769 (9)	0.085742 (8)	0.021190 (2)	0.032298 (5)
	0.019169 (1)	0.008086 (1)	0.067715 (9)	0.029590 (5)	0.087963 (9)	0.021192 (3)	0.032384 (6)
	7^{th} sub-period	8^{th} sub-period	9^{th} sub-period	10^{th} sub-period	11^{th} sub-period	12^{th} sub-period	
Numbers	253	243	159	118	747	668	
Variance	4.26E-04	7.72E-04	3.63E-04	1.24E-03	5.08E-04	3.48E-04	
Skewness	-0.480***	-0.687***	-0.202	-0.838***	-0.520***	0.126	
Kurtosis	3.544*	5.188***	3.263	6.988***	4.741***	3.310*	
AR(2)-GARCH(1,1) RMSE	0.015624 (3)	0.029390 (1)	0.017737 (2)	0.029169 (6)	0.025825 (9)	0.014605 (4)	
	0.015715 (6)	0.029683 (5)	0.017791 (6)	0.028934 (4)	0.025803 (6)	0.014606 (6)	
	0.015628 (4)	0.029866 (9)	0.017785 (4)	0.029198 (7)	0.025804 (7)	0.014605 (4)	
AR(2)-EGARCH(1,1) RMSE	0.015811 (8)	0.029440 (2)	0.017914 (7)	0.028967 (5)	0.025779 (3)	0.014635 (8)	
	0.015557 (1)	0.029621 (4)	0.017940 (8)	0.028857 (2)	0.025758 (1)	0.014635 (8)	
	0.015707 (5)	0.029802 (7)	0.017992 (9)	0.029199 (9)	0.025763 (2)	0.014633 (7)	

Table 7 (cont.). Summary statistics and in-sample RMSEs within the full period and each sub-period

	7th sub-period	8th sub-period	9th sub-period	10th sub-period	11th sub-period	12th sub-period	
AR(2)-TGARCH(1,1) RMSE	0.015811 (8)	0.029598 (3)	0.017732 (1)	0.028853 (1)	0.025807 (8)	0.014595 (1)	
	0.015600 (2)	0.029695 (6)	0.017760 (3)	0.028897 (3)	0.025779 (3)	0.014597 (2)	
	0.015717 (7)	0.029857 (8)	0.017785 (4)	0.029198 (7)	0.025789 (5)	0.014598 (3)	

Notes: The root mean square error (RMSE) estimates of each model with normal distribution is the top row, t-distribution is the middle row, and GED is the bottom row. Ranks are in parentheses. ***, **, and * represent statistical significance at the 0.01, 0.05, and 0.10 levels, respectively.

Table 8 reports the aggregated ranks within each sub-period identified via the ICSS algorithm. Both the AR-EGARCH-*t* and the AR-TGARCH-normal models rank lowest in four of the twelve sub-periods. The last column of Table 8 indicates a score (sum of the numbers multiplied by their corresponding rank in each row). The AR-EGARCH-*t* model exhibits the lowest (46) score and appears to be the most effective for modeling oil futures returns. More specifically, the AR-EGARCH-GED model performs relatively poorly in each of the twelve sub-periods in terms of the model evaluation relative to the superior ICSS AR-EGARCH-GED model for oil futures volatility modeling.

Table 8. In-sample RMSE ranks within each sub-period

Models \ Ranks	1	2	3	4	5	6	7	8	9	Score
AR(2)-GARCH(1,1)	2	2	1	1	0	2	2	1	1	56 (3)
	0	0	2	3	2	5	0	0	0	58 (6)
	0	1	1	4	0	0	3	1	2	68 (8)
AR(2)-EGARCH(1,1)	1	2	1	1	2	2	1	2	0	57 (4)
	4	2	0	2	1	0	0	2	1	46 (1)
	0	2	1	0	1	1	4	1	2	72 (9)
AR(2)-TGARCH(1,1)	4	1	1	1	0	1	0	4	0	51 (2)
	0	3	3	0	2	1	0	1	2	57 (4)
	1	0	2	1	2	1	2	1	2	67 (7)

Notes: The numbers of RMSE ranks of each model with normal distribution is the top row, t-distribution is the middle row, and GED is the bottom row. Score is the sum of the numbers multiplied by their corresponding rank in each row.

Conclusions

In this study, we dive into an investigation of the presence of volatility changes and the heavy-tailed behavior when modeling oil futures market returns. The multiple structural breaks in variance are detected using the Inclán and Tiao (1994) test. The estimated results suggest that the changing oil futures risk can be best captured by the ICSS AR-EGARCH-GED model. The in-sample comparison of the model evaluation shows that the AR-EGARCH-*t* model outperforms over competing models within each sub-period identified using the ICSS algorithm. Our empirical results are provided to illustrate the importance of incorporating both structural breaks and heavy-tailed distributions in oil futures price modeling.

References

1. Aggarwal, R., Inclán, C. and Leal, R. (1999). Volatility in Emerging Markets, *Journal of Financial and Quantitative Analysis*, 34 (1), pp. 33-55.
2. Bracker, K. and Smith, K.L. (1999). Detecting and Modeling Changing Volatility in the Copper Futures Market, *Journal of Futures Markets*, 19 (1), pp. 79-100.
3. Diebold, F.X. (1986). Modeling the Persistence of Conditional Variances: A Comment, *Econometric Reviews*, 5 (1), pp. 51-56.
4. Engle, R.F. and Ng, V.K. (1993). Measuring and Testing the Impact of News on Volatility, *Journal of Finance*, 48 (5), pp. 1749-1778.
5. Ewing, B.T. and Malik, F. (2005). Re-examining the Asymmetric Predictability of Conditional Variances: The Role of Sudden Changes in Variance, *Journal of Banking and Finance*, 29 (10), pp. 2655-2673.
6. Glosten, L.R., Jaganathan, R., and Runkle, D. (1993). On the Relation between the Expected Value and the Volatility of the Nominal Excess Return on Stocks, *Journal of Finance*, 48 (5), pp. 1779-1801.
7. Huang, P.K. (2014). The Stock Index Futures Hedge Ratio with Structural Changes, *Investment Management and Financial Innovations*, 11 (1), pp. 97-113.

8. Inclán, C. and Tiao, G.C. (1994). Use of Cumulative Sums of Squares for Retrospective Detection of Changes of Variance, *Journal of the American Statistical Association*, 89 (427), pp. 913-923.
9. Lamoureux, C.G. and Lastrapes, W.D. (1990). Persistence in Variance, Structural Change, and the GARCH Model, *Journal of Business Economic Statistics*, 8 (2), pp. 225-234.
10. Lastrapes, W.D. (1989). Exchange Rate Volatility and U.S. Monetary Policy: An ARCH Application, *Journal of Money, Credit, and Banking*, 21 (1), pp. 66-77.
11. Malik, F. (2003). Sudden Changes in Variance and Volatility Persistence in Foreign Exchange Markets, *Journal of Multinational Financial Management*, 13 (3), pp. 217-230.
12. Malik, F. and Hassan, S.A. (2004). Modeling Volatility in Sector Index Returns with GARCH Models Using an Iterated Algorithm, *Journal of Economics and Finance*, 28 (2), pp. 211-225.
13. Mansur, I., Cochran, S.J., and Shaffer, D. (2007). Foreign Exchange Volatility Shifts and Futures Hedging: An ICSS-GARCH Approach, *Review of Pacific Basin Financial Markets and Policies*, 10 (3), pp. 349-388.
14. McMillan, D.G. and Wohar, M.E. (2011). Structural Breaks in Volatility: The Case of UK Sector Returns, *Applied Financial Economics*, 21 (15), pp. 1079-1093.
15. Nelson, D.B. (1991). Conditional Heteroskedasticity in Asset Returns: A New Approach, *Econometrica*, 59 (2), pp. 347-370.
16. Wilson, B., Aggarwal, R., and Inclán, C. (1996). Detecting Volatility Changes across the Oil Sector, *Journal of Futures Markets*, 16 (3), pp. 313-330.
17. Zakoïan, J.M. (1994). Threshold Heteroskedastic Models, *Journal of Economic Dynamics and Control*, 18 (5), pp. 931-944.