Quantile Serial Dependence in Crude Oil Markets: Evidence from Improved Quantilogram Analysis with Quantile Wild Bootstrapping

Jen-Je Su[a], Adrian (Wai-Kong) Cheung[b] and Eduardo Roca[a]

[a] Department of Accounting, Finance and Economics, Griffith University, Nathan, Queensland, 4111, Australia
[b] Department of Finance and Banking, School of Economics and Finance, Curtin University, Bentley, Western Australia 6102, Australia

Abstract

We examine the quantile serial dependence in crude oil prices based on the Linton and Whang's (2007) quantile-based portmanteau test which we improved by means of quantile wild bootstrapping. Through Monte Carlo simulation, we find that the quantile wild bootstrap based portmanteau test performs better than the bound testing procedure suggested by Linton and Whang (2007). We apply the improved test to examine the efficiency of two crude oil markets – WTI and Brent. We also examine if the dependence is stable via rolling sample tests. Our results show that both WTI and Brent are serially dependent in all, except the median quantiles. These findings suggest that it may be misleading to examine the efficiency of crude oil markets in terms of mean (or median) returns only. These crude oil markets are relatively more serially dependent in non-median ranges.

Keywords: crude oil prices, efficiency, quantile serial dependence, quantilogram, wild bootstrapping.

JEL Classification Codes: Q40, C12, C15, C22, C50

[1] Introduction

It is well-accepted that the crude oil market is very important as it can significantly affect the performance of the economy and financial markets (see, among others, Hamilton, 2011; Huang, et al, 1996; Park and Ratti, 2008; Lee et al., 2012). Several studies have found statistically significant evidences that an increase in oil price is the key contributing factor behind recessions (see, for examples, Hamilton, 1983; Barsky and Lutz, 2004). Due to its central role in the world economy, it is imperative to understand how the crude oil market works. For example, central banks and other economic policy makers view oil price as one of the key variables in generating macroeconomic projections and in assessing macroeconomic risks.

Ever since the issue of market efficiency was brought to the forefront by the work of Fama in the 1970s, a voluminous amount of studies have been conducted on this issue in different financial and economic markets. The scale of research on this issue in the crude oil market is quite large. As a matter of fact, this efficiency issue has been examined through various lens or approaches -- including, for examples, variance ratio tests (Charles and Darne, 2009), unit root tests (Elder and Serletis, 2008; Maslyuk and Smyth, 2009), time varying long-range dependence (Tabak and Cajueiro, 2007), Hurst exponent dynamics from detrended fluctuation analysis (Alvarez-Ramirez et al., 2008; Wang and Liu, 2010), and neural network (Yu, et al, 2008). However, the focus in the literature has been on the first moment (level or conditional mean) or second moment (volatility) of oil prices/returns and little is known beyond the first or second moments. In addition, the research on predictability in this market has yielded mixed results (Alquist, et al, 2013). Thus, there is a need for further research on market efficiency in the crude oil market with a new perspective/approach. Our paper, therefore, addresses this gap in the literature.

Our approach is to look at the efficiency issue of the crude oil markets in terms of serial dependence through the application of Linton and Whang's (2007) quantile-based Portmanteau test. This approach is special as it can be used to measure serial dependence in *different quantiles* of the crude oil prices distributions. In other words, our concern is beyond the oil price dynamics at the middle as other parts (upper and lower tails) of the crude oil price distribution are also considered. This is a very important issue since oil prices have been shown to exhibit fat tails (Nordhaus, 2007; Trolle and Schwartz, 2010). To the best of our knowledge, this paper is the first in the literature to examine the issue by means of Linton and Whang's test.

In addition, this paper also contributes to the econometrics literature. Instead of employing the inference strategy suggested in Linton and Whang (2007), which at times lead to inclusive results, we conduct the Linton and Whang's (2007) test by means of the quantile wild bootstrapping of Feng et al (2011). We show, via simulations, that the bootstrap-based inference is accurate in size without compromising in testing power and it can effectively avoid inconclusive outcomes.

As an overview, first, our results show that by means of wild bootstrapping, we were able to improve the finite sample properties of the Linton and Whang's (2007) quantile-based Portmanteau test. By applying this improved test to the analysis of the efficiency of WTI and Brent in crude oil markets, we have found that both WTI and Brent are serially dependent in all but the median quantiles. Interestingly, our rolling sub-period results also reveal that the Brent oil market is with higher degree of serial dependence than the WTI market in the lower quantiles while the opposite is true in the higher quantiles. These findings suggest that it may be misleading to examine the efficiency issue of crude oil markets in terms of mean (or median) oil returns only. The result also implies that these crude oil markets may be predictable in non-median ranges. The rest of the paper is organised as follows. Section 2 discusses the methodology and Section 3 presents the empirical results. Section 4 concludes the study.

[2] Methodology

In this section, we discuss the quantilogram and the quantile portmanteau test proposed by Linton and Whang (2007). Specifically, we explain the strengths and limitations of the associated inference of these statistics and how we improve the inference through quantile wild bootstrapping introduced in Feng et al (2011). We present the results of the Monte Carlo simulation which demonstrate the improvement.

2.1. Quantilogram and quantile portmanteau test

Suppose that Y_1, Y_2, … are random variables drawn from a stationary process whose marginal distribution has quantiles μ_q for $0<q<1$. Linton and Whang (2007) define the quantilogram for any q as

$$\rho_q(k) = \frac{E[\psi_q(Y_t - \mu_q)\,\psi_q(Y_{t+k} - \mu_q)]}{E[\psi_q^2(Y_t - \mu_q)]}, \text{ k=1, 2, …} \quad (1)$$

where $\psi_q(x) = q - 1(x<0)$ denotes the check function. Since the quantilogram $\rho_q(k)$ considers the dynamic association in terms of the direction of deviation from a given quantile, it can be used to measure the quantile serial dependence of the stochastic process

$\{Y_t\}_{t=1}^{\infty}$. The issue is whether the past information set of Y_t can or cannot be used to improve the prediction about whether Y_t will be above or below a given quantile, μ_q. Under the null hypothesis of no serial dependence, $\rho_q(k) = 0$ for all k. Alternatively, if $\rho_q(k) \neq 0$ for some k, $\{Y_t\}_{t=1}^{\infty}$ is serially dependent.

Let $\hat{\mu}_q$ be the sample quantile obtained as a solution of the minimization problem: $\min_{\mu \in R} \sum_{t=1}^{T} \Theta_q(Y_t - \mu)$ where $\Theta_q(x) = x[q - 1(x < 0)]$. The sample counterpart of $\rho_q(k)$ can be computed as follows:

$$\hat{\rho}_q(k) = \frac{\sum_{t=1}^{T-k} \psi_q(Y_t - \hat{\mu}_q)\psi_q(Y_{t+k} - \hat{\mu}_q)}{\sqrt{\sum_{t=1}^{T-k} \psi_q^2(Y_t - \hat{\mu}_q)}\sqrt{\sum_{t=1}^{T-k} \psi_q^2(Y_{t+k} - \hat{\mu}_q)}}, \text{ k=1,2,..., T-1.} \quad (2)$$

Since $\hat{\rho}_q(k)$ is constructed as a sample autocorrelation of the check function (i.e. the sample correlation of $\psi_q(Y_t - \hat{\mu}_q)$ and $\psi_q(Y_{t+k} - \hat{\mu}_q)$), it should lie between 0 and 1 for any k and q. To test the null hypothesis of no directional predictability at *q* up to p lags (i.e. $\rho_q(k) = 0$ for k=1,..., p), Linton and Whang (2009) suggest a quantile version of Box-Ljung Q test (henceforth, QQ):

$$QQ_q(p) = T(T+2)\sum_{k=1}^{p} \hat{\rho}_q^2(k)/(T-k). \quad (3)$$

As the usual portmanteau Q test, the interpretation of the $QQ_q(p)$ test is straightforward, if the null hypothesis cannot be rejected, there exhibits insufficient evidence against serial dependence (at q); instead, if the null hypothesis is rejected, the underlying series is serially dependent. The inference of $\hat{\rho}_q(k)$ and $QQ_q(p)$ is, however, not straightforward. Specifically, under the null hypothesis, the asymptotic distribution of $[\hat{\rho}_q(1),...,\hat{\rho}_q(p)]'$, as shown by Theorem 2 in Linton and Whang (2007), is

$$\sqrt{T}\begin{bmatrix}\hat{\rho}_q(1)\\ \vdots \\ \hat{\rho}_q(p)\end{bmatrix} \Rightarrow N(0,V_q). \tag{4}$$

Here, V_q is a $p\times p$ asymptotic variance-covariance matrix which, in general, depends on the underlying volatility process of Y_t. Since $QQ_q(p)$ is a function of $\hat{\rho}_q(1)$,..., $\hat{\rho}_q(p)$, $QQ_q(p)$ is not asymptotically valid. To avoid the necessity of specifying a volatility model, Linton and Whang (2007) derive the lower and upper bounds of $V_{q,kk}$, the kth diagonal component of $V_q^{(p)}$, $1\le V_{q,kk}\le 1+\overline{v}_q$, and the (j, k) off-diagonal component, $|V_{q,jk}|\le \overline{v}_q$, where $\overline{v}_q=[\max(q,1-q)]^2/q(1-q)$.

Under the null hypothesis, when the upper bound is considered, the $(1-\alpha)\%$ confidence interval of $\hat{\rho}_q(k)$ can be constructed as $CI_1=\left(-z_{\alpha/2}\sqrt{(1+\overline{v}_q)/T},\ z_{\alpha/2}\sqrt{(1+\overline{v}_q)/T}\right)$. We note that, since $\overline{v}_q$ increases without limit as $q\to 0,1$, the confidence interval can be very wide when extreme quantiles are considered. In some special circumstances (e.g. conditions that satisfy equation (6) in Linton and Whang (2007)), V_q is the identity matrix and so the lower bound can be applied, the confidence interval shrinks to be $CI_2=\left(-z_{\alpha/2}\sqrt{1/T},z_{\alpha/2}\sqrt{1/T}\right)$. Linton and Whang (2007) call the larger band CI_1 conservative and the smaller band CI_2 liberal. Correspondingly, the decision rule of applying the $QQ_q(p)$ test can be either conservative or liberal: according to the conservative rule, the null hypothesis is rejected if $QQ_q(p)>(1+p\overline{v}_q)\chi^2_\alpha(p)$ and for the liberal rule if $QQ_q(p) > \chi^2_\alpha(p)$. Such an inference setup is indeed very novel in circumventing the complication of modelling the volatility process. Thus, with this setup, the QQ test is valid even in the case where the first or even the second moment is infinite (e.g., in the case of α-stable process). However, in practice, the user needs to decide which rule – conservative or liberal – should be employed. The test would gain power at the cost of potential size distortion when the liberal rule is considered. On the other hand, with the conservative rule, the test can be very conservative for even moderate p (specially, when *q* is close 0 or 1) and has no power to reject the null hypothesis even if the tested process is indeed serially dependent. On the other hand, if both rules are applied simultaneously, the testing result might turn out to be inconclusive if the test statistic locates between the two – liberal and conservative – critical values.

2.2. Quantile wild bootstrapping

In this paper, we suggest approximating the distribution of $\hat{\rho}_q(k)$ and $QQ_q(p)$ via quantile wild bootstrapping (QWB). Specifically, we adopt and modify the bootstrapping procedure of Feng et al (2011) as follows.

(1) Obtain $\hat{e}_t = Y_t - \hat{\mu}_q$. Form a bootstrap sample of T observations $Y_t^* = \hat{\mu}_q + \omega_t |\hat{e}_t|$ where the weights ω_t is generated as $\omega_t = \begin{cases} 2(1-q) \text{ with probability } p = 1-q \\ -2q \text{ with probability } p = q. \end{cases}$

(2) Compute $\hat{\rho}_q(k)$ and $QQ_q(p)$ based on $\{Y_t^*\}_{t=1}^T$ and label the associated statistics: $\hat{\rho}_q^*(k)$ and $QQ_q^*(p)$.

(3) Repeat (1) and (2) B times to form bootstrap distribution of $\hat{\rho}_q(k)$ and $QQ_q(p)$:

$\{\hat{\rho}^*_{q,b}(k)\}_{b=1}^{B}$ and $\{QQ^*_{q,b}(p)\}_{b=1}^{B}$, respectively .

For the individual quantilogram, a $100(1-\alpha)\%$ confidence interval for $\rho_q(k)$ can be obtained as $\left[\hat{\rho}_q(k)-T^{-1/2}c^*_{1k,\alpha},\hat{\rho}_q(k)+T^{-1/2}c^*_{2k,\alpha}\right]$ where $\left(c^*_{1k,\alpha},c^*_{2k,\alpha}\right)$ are from the bootstrap distribution of $\left\{\left(\hat{\rho}^*_{q,b}(k)-\hat{\rho}_q(k)\right)\right\}_{b=1}^{B}$ such that $\Pr(c^*_{1k,\alpha}\le T^{1/2}\left(\hat{\rho}^*_{q,b}(k)-\hat{\rho}_q(k)\right)\le c^*_{2k,\alpha})=1-\alpha$. For the test of no directional predictability with the $QQ_q(p)$ test, a critical value, $c^*_{QQ,\alpha}$, for a significance level α is given by $c^*_{QQ,\alpha}=\inf\left\{c:\Pr\left(QQ^*_{q,b}(p)<c\right)\ge 1-\alpha\right\}$, the $(1-\alpha)100\%$ percentile of $\{QQ^*_{q,b}(p)\}_{b=1}^{B}$.

2.3. Monte Carlo simulation results

The simulation design is modified from Linton and Whang (2009). Specifically, we consider three models: (A) I.I.D. Normal: $Y_t\sim N\left(0,1\right)$; (B) GARCH: $Y_t\sim\varepsilon_t\sigma_t$ with $\sigma_t^2=0.02+0.9\sigma_{t-1}^2+0.05Y_{t-1}^2$ where $\varepsilon_t\sim N\left(0,1\right)$; (C) Threshold-GARCH: $Y_t\sim\varepsilon_t\sigma_t$ with σ_t^2 $=0.2+0.9\sigma_{t-1}^2+0.06Y_{t-1}^2+0.03Y_{t-1}^2 1(Y_{t-1}<0)$. Models (B) and (C) are set to reflect the estimated GARCH and Threshold-GARCH estimation of the crude oil returns considered in the next section. Theoretically, Model (A) is completely under the null hypothesis of no directional predictability for *all* quantiles, while (B) and (C) are under the null for the *median (q=0.5)* and the alternative when other quantiles (q≠0.5) are considered. We choose sample size T=500, 1000, and 5000 and quantile q=0.1, 0.3, 0.5, 0.7, and 0.9. The Monte Carlo simulation is conducted via the warp-speed method of Giacomini et al (2013) with 10000 replications and done by GAUSS. We report the simulation results of the 5% nominal level of $QQ_q(p)$ in Table 1. For the purpose of comparison (in terms of size control and testing power), we report simulation results of $QQ_q(p)$ via both bootstrapping (QWB) and the liberal rule. We do not report results based on the conservative rule as the rule generally is with very low power, especially when lager lags are considered.

Table 1 shows that in the cases when the null is true (Model (A) and q=0.5 in Models (B) and (C)), the bootstrapping inference is corrected sized (i.e. the simulated rejection rate is close to the nominal level of significance at 5%) regardless lag (p), sample size (T), and quantile (q) (Model (A)). In contrast, the testing result can become rather over-sized when the liberal decision rule is applied, particularly when p is large (p=50 or 100) and T is moderate (T=500 or 1000). For example, in Model (A), while the rejection rate at q=0.5 for the bootstrap test is around 0.05 regardless p and T, the rejection rate for the liberal test can range from 0.051 (p=1 and T=1000) to 0.240 (p=100, T=500). Therefore, the bootstrapping approach is advantageous over the liberal rule in controlling for the Type I error when the sample size is not very large.

In the cases under the alternative (Model (B) and (C) with $q \neq 0.5$), Table 1 shows that, in most cases, the bootstrap test exhibits competitive power compared with the liberal test. For example, in all cases with p=1 and 5, the power of the two testing strategies is indistinguishable. We note that for those cases that the liberal test is significantly more powerful than the bootstrap test, the liberal test also suffers considerable size distortion while the bootstrap test does not – say, for example, when p=50, 100 and T=500, 1000. Overall, our simulation results support the use of QWB.

[3] Empirical analysis of crude oil markets

We apply the proposed QWB-based quantilogram in analysing the quantile serial dependence in two major crude oil markets: WTI and Brent. We are the first to apply the quantilogram analysis in energy markets.

3.1. Data

For the two crude oil markets, we collect daily spot prices from Energy Information Administration website.[1] Both prices collected end on December 1, 2014 but with different starting dates (due to data availability): WTI is from January 2, 1986 with 7,207 observations and Brent is from May 20, 1987 with 6,923 observations. We plot the WTI and Brent crude oil prices & returns (log-returns) in Figure 1. In general, the two price series move closely to each other and both returns are very volatile. We also present descriptive statistics for the two oil returns in Table 2. In summary, the two returns are close to zero and both are volatile, left-skewed and leptokurtic.

3.2. Empirical results

We first report the full-sample results of the quantilogram and the corresponding quantile portmanteau test for WTI and Brent returns in Figure 2 and Table 3. We consider the cases with lags up to 100 trading days at five various quantiles (q=0.1, 0.3, 0.5, 0.7, 0.9). We also show the QWB-based 95% confidence intervals (centred to zero) for the quantilogram and the QWB-based 5% critical values (Figure 2), and the p-values for the quantile-portmanteau test (Table 3). The bootstrapping is performed with 1000 replications.

[1] http://www.eia.gov/

Both WTI and Brent returns, as shown in Table 3, seem to exhibit evidence of serial dependence except for the case with q=0.5., The null hypothesis of no serial dependence is rejected at 5% level in nearly all cases when $q \neq 0.5$ but not otherwise (i.e., when *q*=0.5). The autocorrelation at q=0.5 is evidently non-existent for the WTI returns and while there is some evidence of positive median autocorrelation for Brent, the evidence is not strong and the QQ test fails to reject the null of no serial correlation up to 100 lags. Therefore, overall, there is not much evidence of dependence in the median. This result is consistent with the well-documented fact that the conventional autocorrelation of asset/commodity returns is often zero (Pagan, 1996), a phenomenon implied by the market efficiency hypothesis. In other words, autocorrelation is expected to be zero because it is limited by arbitrage.

In the lower quantiles, when q=0.1 in particular, there are many cases of individually significant positive serial dependence and the dependence appears to be strong and persistent. Comparing the two markets, as shown in Figure 2 and Table 3, WTI seems to be more dependent than Brent at q=0.1 while Brent is somewhat more dependent than WTI at q=0.3. . Thus, in general, when there are large losses in one period, the chance of having large losses in the next few periods is high (higher than 10% in the case of q=0.1, unconditionally). Similarly, both WTI and Brent are serially dependent for the higher quantiles. There are significant long-lasting, positive serial-correlations – implying that when there are large gains in one period, the chance of having large gains in the next few periods is also high. These results indicate presence of volatility clustering that is, as first noted in Mandelbrot (1963), recognized as a stylized property present in many speculative price time series. This volatility clustering effect has given rise to the development of stochastic models – GARCH models and stochastic volatility models are intended primarily to model this phenomenon in oil returns. Economists have identified some economic mechanisms that could lead to volatility clustering effect. They include herd behaviour of market participants (Lux and Marchesi, 2000), heterogeneous arrival rates of information (Andersen and Bollerslev, 2007), and leverage (Thurner, Farmer and Geanakoplos, 2012), among others.

To examine if the dependence is stable across time, we also run rolling quantile portmanteau test ($\hat{Q}Q_q(p)$) with a four-year window (each with 1,008 daily returns) moving up by three months (63 observations). For WTI, there are 95 rolling results and for Brent 90 results. We report the results using p=50 at the 5% significance level with various q's in Figure 3 and summarize the rejection percentage in Table 4. As shown in Table 4, for both oil returns, there are significantly more sub-periods with rejection in the lower and upper quantiles than in the middle. Interestingly, in the case of Brent (but not WTI), it appears that the market is more dependent in the lower quantiles than in the higher quantiles. Specifically, for Brent, at q=0.1 (0.3) there is 71.11% (63.33%) rejection among the examined sub-periods; in contrast, the rejection rate is 54.44% (21.22%) at q=0.9 (0.1). This implies an asymmetric dependence feature in the Brent market. Thus, large losses are somewhat more often coming after large losses than large gains arising after large gains. Moreover, Brent appears to be somewhat more dependent than WTI in the lower quantiles while it is opposite in the higher quantiles. As is well known in the crude oil literature, there are at least two reasons why the WTI and Brent markets and their prices are behave differently. First, the inland U.S. WTI and the seaborne Brent crude oil have been traded in partly segregated markets (see, for example, Büyükşahin et al, 2013). Second, infrastructure constraints in Cushing, Oklahoma, have historically influenced the price differential between the WTI crude and the Brent crude oil trades (Fattouh, 2010; Borenstein and Kellogg, 2012).[2] This effectively makes Cushing oil stocks insulate the WTI market from the price pull stemming from strong world demand, suggesting that WTI is more sensitive to US conditions while Brent is more sensitive to the world conditions. However, of course, how these conditions are actually transmitted into the asymmetric effects that we found at lower and higher quantiles across the two markets, is not clear and deserves further investigation.

We also observe from Figure 3 that the dependent sub-periods at q=0.1 and q=0.9 for both oil markets (WTI, especially) tend to concentrate on the first 13/14 sampling years (around 1986 to 1999) and the last 10/11 years (roughly, 2003/2004 to 2014). A closer look at Figure 1, especially the return series of WTI and Brent, reveals that there were many large (positive or negative) returns during the two periods. In addition, these large returns tend to cluster with each other, confirming the volatility clustering effect depicted in Table 3. By browsing the chronology of economic or political events in the world crude oil market, it is not difficult to understand why there were so many large positive and negative returns in these two periods.[3]

[4] Conclusion

We examined the quantile serial dependence of crude oil prices based on an improved version of the Linton and Whang's (2007) quantile-based portmanteau test. We improved the test by means of quantile wild bootstrapping. Through Monte Carlo simulation, we

[2] Several new crude transportation projects came online in early 2013, including pipelines and crude-by-rail terminals. This new infrastructure helped clear transportation bottlenecks in U.S. Midcontinent, particularly around Cushing, Oklahoma.

[3] The chronology is available at the links below:
https://en.wikipedia.org/wiki/Chronology_of_world_oil_market_events;
https://en.wikipedia.org/wiki/2001_world_oil_market_chronology;
https://en.wikipedia.org/wiki/World_oil_market_chronology_from_2003.

found that the quantile wild bootstrap based portmanteau test corrects the over-sizing problem of the standard portmanteau test. We applied the improved test to examine the efficiency of two crude oil markets – WTI and Brent. Our results showed that both WTI and Brent are dependent in all, except the median quantiles. Interestingly, our rolling subsample results showed that the Brent oil market tends to be more dependent than the WTI market in the lower quantiles while the opposite is true in the higher quantiles. These findings suggest that it may be misleading to examine the efficiency of crude oil markets in terms of mean (or medium) returns only. These crude oil markets are relatively more dependent in non-median ranges.

Reference

Alquist, R., Kilian, L. and Vigfusson, R.J. (2013) Forecasting the price of oil, in: G. Elliott and A. Timmermann (eds*.), Handbook of Economic Forecasting*, 2, Amsterdam: North-Holland, 427-507.

Alvarez-Ramirez, J., Alvarez, J. and Rodriguez, E. (2008) Short-term predictability of crude oil markets: A detrended fluctuation analysis approach. *Energy Economics*, 30, 2645–2656.

Andersen, T., and Bollerslev, T. (1997) Heterogeneous information arrivals and returns volatility dynamics, *Journal of Finance*, 52, 975–1005.

Barsky, R. B. and Kilian L. (2004) Oil and the macroeconomy since the 1970s, *The Journal of Economic Perspectives*, 18, 115-134.

Borenstein, S. and Kellogg, R. (2012) The incidence of an oil glut: who benefits from cheap crude oil in the midwest? *NBER Working Paper No*. 18127, June.

Büyükşahin, B., Lee, T.K., Moser, J.T. and Robe, M.A. (2013) Physical markets, paper markets and the WTI-Brent spread, *Energy Journal*, 34, 129–51.

Charles, A. and Darne, O. (2009) The efficiency of the crude oil markets: Evidence from variance ratio tests, *Energy Policy*, 37, 4267–4272.

Elder, J. and Serletis, A. (2008) Long memory in energy futures prices, *Review of Financial Economics*, 17, 146-155.

Fama, E.F. (1970) Efficient capital markets: A review of theory and empirical work, *The Journal of Finance*, 25, 383-417.

Fattouh, B. (2010) The dynamics of crude oil price differentials. *Energy Economics* 32: 334–342.

Fen, X., He, X. and Hu, J. (2011) Wild bootstrap for quantile regression, *Biometrika,* 98, 995-999.

Giacomini, R., Politis, D.N. and White, H. (2013) A warp-speed method for conducting Monte Carlo experiment in solving bootstrap estimator, *Econometric Theory*, 29, 567-589.

Hamilton, J. D. (1983) Oil and the macroeconomy since World War II*, The Journal of Political Economy*, 91, 228-248.

Hamilton, J.D. (2011) Nonlinearities and the macroeconomic effects of oil prices, *Macroeconomic Dynamic*, 364-378.

Huang, R., Masulis, R. and Stoll, H. (1996) Energy shocks and financial markets, *Journal of Futures Markets*, 16, 1-27.

Lee, B.J., Yang, C.H. and Huang, B.H. (2012) Oil price movements and stock market revisited: a case of sector stock price indexes in the G7 countries, *Energy Economics*, 34, 1284–1300.

Linton, O. and Whang, Y.-J. (2007) The quantilogram: with an application to evaluating directional predictability, *Journal of Econometrics*, 141, 250-282.

Lux, T. and Marchesi, M. (2000) Volatility clustering in financial markets : a micro simulation of interacting agents, *International Journal of Theoretical and Applied Finance*, 3, 675–702.

Maslyuk, S. and Smyth, R. (2009) Unit root properties of crude oil spot and futures prices, *Energy Policy*, 31, 109–118.

Nordhaus, W., (2007) Who's afraid of a big bad oil shock? *Brookings Papers on Economic Activity*, 2007:2.

Pagan, A. (1996) The econometrics of financial markets. *Journal of Empirical Finance*, 3:15-102.

Park, J. and Ratti, R. A. (2008) Oil price shocks and the stock markets in the U.S. and 13European countries, *Energy Economics*, 30(5), 2587-2608.

Tabak, B.M. and Cajueiro, D.O. (2007) Are the crude oil markets becoming weakly efficient over time? A test for time-varying long-range dependence in prices and volatility, *Energy Economics*, 29, 28–36.

Trolle, A. B. and Schwartz, E. S., (2010) Variance risk premia in energy commodities. *The Journal of Derivatives* 17.3, pp. 15–32

Thurner, S., Farmer D.J., and Geanakoplos, J. (2012) Leverage causes fat tails and clustered volatility. *Quantitative Finance*, 12, 695–707.

Wang, Y.-D. and Liu, L. (2010) Is WTI crude oil market becoming weakly efficient over time? New evidence from multiscale analysis based on detrended fluctuation analysis, *Energy Economics*, 32, 987–992.

Yu, L. A., Wang, S.-Y. and Lai, K. K., (2008) Forecasting crude oil price with an EMD-based neural network ensemble learning paradigm, *Energy Economics*, 30, 2623–2635.

Table 1. (Simulation) Shows empirical rejection frequency of the quantile portmanteau test (Equation (3)) based on quantile wild bootstrapped (QWB) and liberal critical values at the 5% level against lags p=1, 5, 50 and 100. Design Model (A): IID-N(0,1), Model (B) GARCH and Model (C): Threshold-GARCH (described in Section 2.3) with sample size T=500, 1000 and 5000.

(A) IID-N(0,1)

		T=500		T=1000		T=5000	
Quantile (q)	**Lag(p)**	**QWB**	**Liberal**	**QWB**	**Liberal**	**QWB**	**Liberal**
0.1	**1**	0.051	0.039	0.048	0.046	0.049	0.048
	5	0.051	0.046	0.047	0.047	0.050	0.050
	50	0.053	0.094	0.051	0.070	0.048	0.054
	100	0.054	0.205	0.051	0.119	0.048	0.061
0.3	**1**	0.049	0.055	0.051	0.050	0.051	0.051
	5	0.053	0.055	0.052	0.052	0.046	0.050
	50	0.049	0.105	0.051	0.076	0.050	0.054
	100	0.049	0.228	0.052	0.126	0.048	0.061
0.5	**1**	0.048	0.048	0.044	0.051	0.053	0.053
	5	0.052	0.054	0.048	0.052	0.052	0.052
	50	0.050	0.111	0.050	0.077	0.053	0.057
	100	0.051	0.240	0.048	0.122	0.051	0.064
0.7	**1**	0.048	0.054	0.052	0.053	0.051	0.049
	5	0.051	0.052	0.049	0.052	0.054	0.051
	50	0.049	0.106	0.052	0.077	0.051	0.056
	100	0.050	0.227	0.049	0.122	0.050	0.062
0.9	**1**	0.049	0.047	0.050	0.054	0.053	0.050
	5	0.052	0.047	0.050	0.049	0.048	0.049
	50	0.052	0.094	0.051	0.070	0.054	0.056
	100	0.050	0.209	0.055	0.116	0.051	0.061

(B) GARCH

		T=500		T=1000		T=5000	
Quantile (q)	**Lag(p)**	**QWB**	**Liberal**	**QWB**	**Liberal**	**QWB**	**Liberal**
0.1	**1**	0.119	0.114	0.150	0.147	0.452	0.441
	5	0.190	0.187	0.282	0.285	0.808	0.818
	50	0.262	0.310	0.415	0.440	0.953	0.955
	100	0.224	0.387	0.354	0.438	0.911	0.920
0.3	**1**	0.056	0.056	0.057	0.055	0.076	0.082
	5	0.059	0.063	0.073	0.070	0.118	0.127
	50	0.079	0.112	0.075	0.101	0.166	0.177
	100	0.089	0.239	0.079	0.151	0.142	0.162
0.5	**1**	0.047	0.047	0.045	0.051	0.051	0.051
	5	0.052	0.055	0.055	0.053	0.051	0.051
	50	0.062	0.102	0.055	0.074	0.058	0.058
	100	0.079	0.231	0.068	0.137	0.057	0.064
0.7	**1**	0.048	0.059	0.057	0.057	0.076	0.086
	5	0.053	0.056	0.060	0.063	0.111	0.122
	50	0.074	0.120	0.074	0.100	0.161	0.173
	100	0.089	0.243	0.075	0.146	0.140	0.166
0.9	**1**	0.091	0.090	0.162	0.176	0.468	0.473
	5	0.181	0.179	0.271	0.277	0.812	0.823
	50	0.263	0.311	0.404	0.434	0.948	0.951
	100	0.231	0.388	0.356	0.438	0.915	0.921

(C) Threshold-GARCH

		T=500		T=1000		T=5000	
Quantile (q)	**k**	**QWB**	**Liberal**	**QWB**	**Liberal**	**QWB**	**Liberal**
0.1	**1**	0.198	0.195	0.286	0.286	0.832	0.831
	5	0.342	0.333	0.552	0.550	0.994	0.994
	50	0.367	0.439	0.626	0.660	0.998	0.999
	100	0.280	0.494	0.530	0.628	0.996	0.997
0.3	**1**	0.065	0.068	0.079	0.079	0.177	0.178
	5	0.079	0.084	0.105	0.107	0.350	0.335
	50	0.084	0.145	0.120	0.156	0.443	0.450
	100	0.073	0.268	0.091	0.190	0.345	0.373
0.5	**1**	0.051	0.051	0.049	0.049	0.049	0.052
	5	0.053	0.055	0.050	0.052	0.052	0.053
	50	0.048	0.103	0.050	0.076	0.048	0.055
	100	0.050	0.231	0.050	0.126	0.049	0.061
0.7	**1**	0.055	0.063	0.064	0.066	0.106	0.107
	5	0.067	0.069	0.078	0.076	0.154	0.158
	50	0.065	0.120	0.081	0.106	0.187	0.199
	100	0.057	0.242	0.067	0.151	0.147	0.170
0.9	**1**	0.125	0.120	0.248	0.250	0.683	0.692
	5	0.265	0.259	0.421	0.422	0.961	0.962
	50	0.293	0.366	0.489	0.529	0.989	0.990
	100	0.222	0.431	0.400	0.512	0.975	0.978

Table 2. Descriptive statistics of daily crude oil returns.

	WTI	Brent
Mean	0.008642	0.017231
Median	0.056101	0.021148
Maximum	19.15065	18.12974
Minimum	-40.63958	-36.12144
Std. Dev.	2.505625	2.262667
Skewness	-0.794085	-0.678947
Kurtosis	18.42884	18.61910
Jarque-Bera (p-value)	72241.62 (0.000)	70903.07 (0.000)
Sample	Jan 2, 1986 - Dec 1, 2014	May 20, 1987 - Dec 1, 2014
Observations	7207	6923

Table 3. Test for no quantile serial correlation of daily crude oil returns based on the quantile portmanteau test (Equation (3)) against lags=1, 10, 50 and 100 with quantiles q=0.1, 0.3, 0.5, 0.7 and 0.9. Quantile wild bootstrapped p-values reported in the parentheses.

WTI

Lag/Quantile	q=0.1	q=0.3	q=0.5	q=0.7	q=0.9
P=1	59.24	6.743	1.570	0.338	12.36
	(0.000)	(0.007)	(0.213)	(0.555)	(0.001)
P=10	291.4	25.42	9.272	40.67	194.0
	(0.000)	(0.004)	(0.501)	(0.000)	(0.000)
P=50	962.2	104.0	47.18	239.0	911.4
	(0.000)	(0.000)	(0.599)	(0.000)	(0.000)
P=100	1444	160.2	88.38	384.7	1379
	(0.000)	(0.000)	(0.801)	(0.000)	(0.000)

Brent

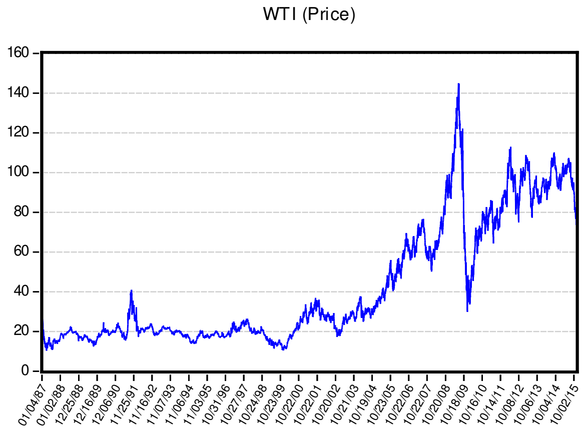
Lag/Quantile	q=0.1	q=0.3	q=0.5	q=0.7	q=0.9
P=1	48.27	22.42	3.994	2.343	13.87
	(0.000)	(0.000)	(0.050)	(0.126)	(0.001)
P=10	214.4	64.75	11.43	64.61	128.7
	(0.000)	(0.000)	(0.338)	(0.000)	(0.000)
P=50	566.7	158.4	54.7	237.7	687.7
	(0.000)	(0.000)	(0.306)	(0.000)	(0.000)
P=100	841.6	236.0	117.5	407.4	1223
	(0.000)	(0.000)	(0.125)	(0.000)	(0.000)

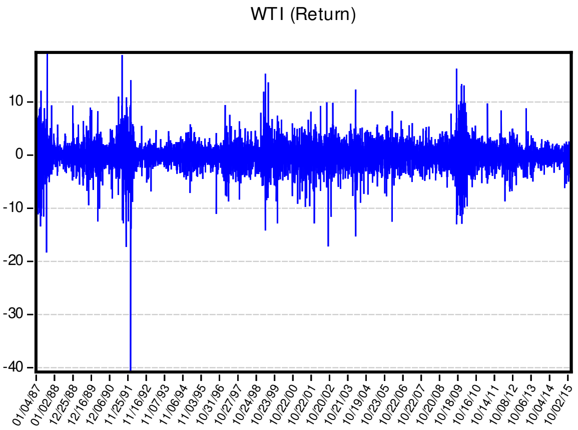
Table 4. (Rolling subsample) Rejection frequency of no quantile serial correlation (up to 50 lags) for daily crude oil returns based on the quantile portmanteau test (Equation (3)) at the 5% significance level (bootstrapped) of 5-year rolling subsamples with 3-month shifts at quantiles q=0.1, 0.3, 0.5, 0.7 and 0.9.

Quantile	WTI	Brent
q=0.1	67.36%	71.11%
q=0.3	22.10%	63.33%
q=0.5	7.37%	20.00%
q=0.7	35.79%	21.11%
q=0.9	63.16%	54.44%

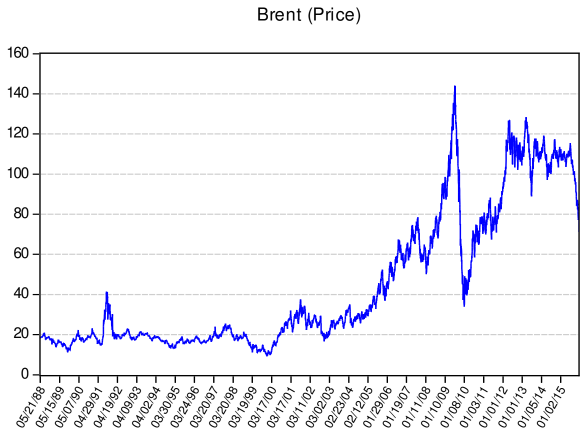
Figure 1: Daily WTI and Brent crude oil prices (left column) and returns (right column)

WTI





Brent



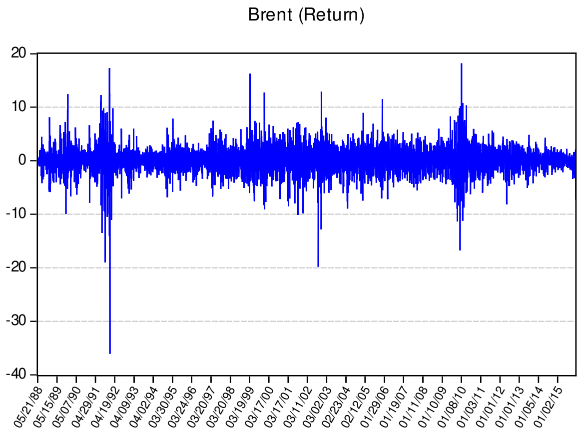
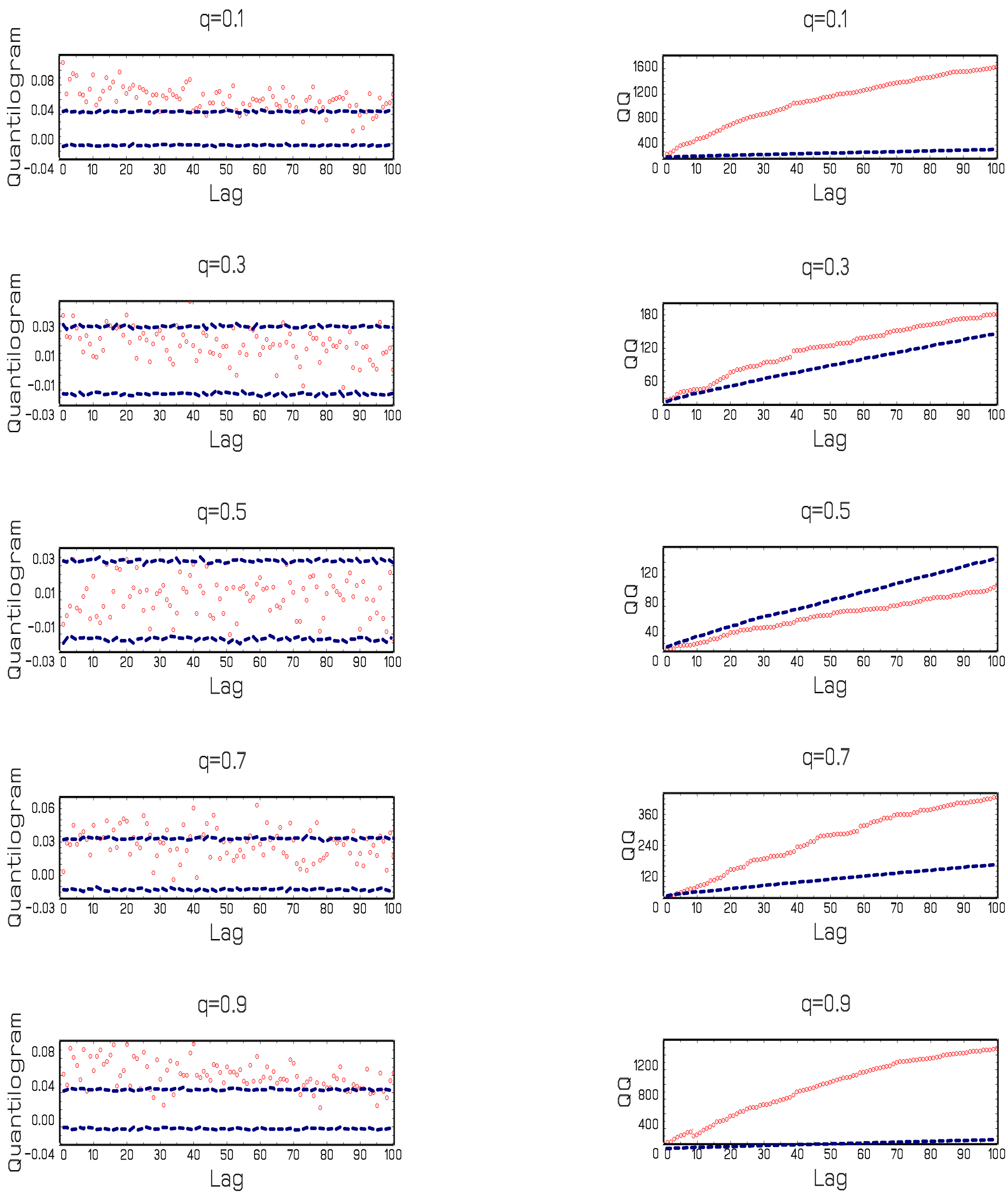


Figure 2. Quantilogram and quantile portmanteau test of daily crude oil returns at quantiles q=0.1, 0.3, 0.5, 0.7 and 0.9 for lags=1, …, 100.

(A) WTI: (Left column) Red dots shown the values of quantilogram, blue dash lines represent the 95% confidence intervals cantered at zero. (Right column) Red dots shown the values of the QQ test, blue dash lines give the bootstrapped 5% critical values.



(B) Brent: (Left column) Red dots shown the values of quantilogram, blue dash lines represent the 95% confidence intervals cantered at zero. (Right column) Red dots shown the values of the QQ test, blue dash lines give the bootstrapped 5% critical values.

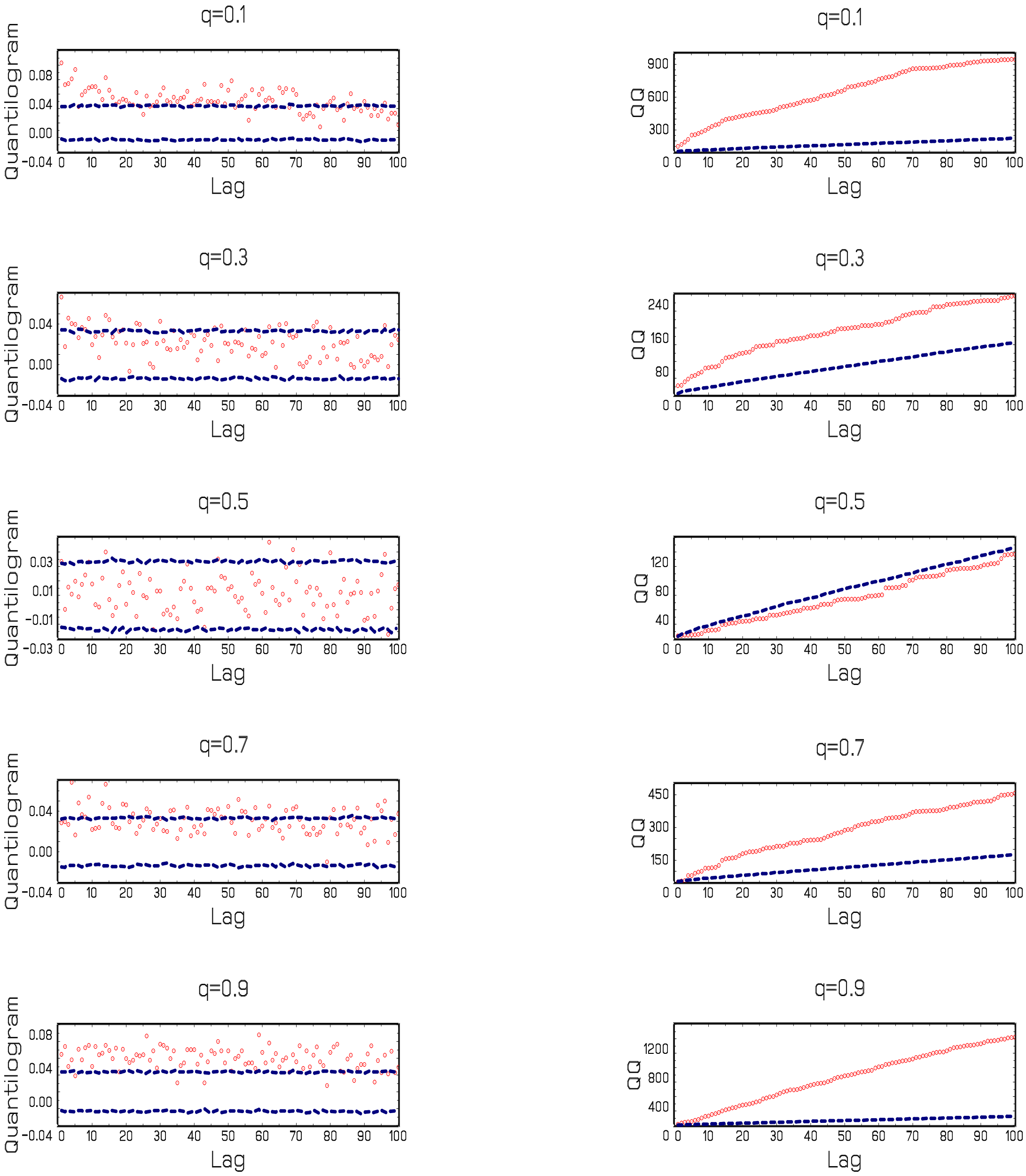


Figure 3. Rolling quantile portmanteau test of daily crude oil returns at quantiles q=0.1, 0.3, 0.5, 0.7 and 0.9 for lag=50.

WTI: Blue lines show the p-value of the quantile portmanteau test of 95 overlapped 5-year rolling subsamples with 3-month shifts. Red dashed lines give the 5% significance level. Shade areas cover subsample periods with rejection of zero quantilogram (i.e. p<0.05).

Brent: Blue lines show the p-value of the quantile portmanteau test of 90 overlapped 5-year rolling subsamples with 3-month shifts. Red dashed lines give the 5% significance level. Shade areas cover subsample periods with rejection of zero quantilogram (i.e. p<0.05).