

# The efficiency of the crude oil markets: Evidence from variance ratio tests

**Amélie CHARLES<sup>\*</sup>**

*Audencia Nantes, School of Management*

**Olivier DARNÉ<sup>†</sup>**

*LEMNA, University of Nantes*

## Abstract

This study examines the random walk hypothesis for the crude oil markets, using daily data over the period 1982–2008. The weak-form efficient market hypothesis for two oil crude markets (UK Brent and US West Texas Intermediate) is tested with non-parametric variance ratio tests developed by Wright (2000) and Belaire-Franch and Contreras (2004) as well as the wild-bootstrap variance ratio tests suggested by Kim (2006). We find that the Brent crude oil market is weak-form efficiency while the WTI crude oil market seems to be inefficiency on the 1994–2008 sub-period, suggesting that the deregulation have not improved the efficiency on the WTI crude oil market in the sense of making returns less predictable.

**Keywords:** Crude oil markets; market efficiency; random walk hypothesis; variance ratio test.

**JEL Classification:** Q40; C14.

---

<sup>\*</sup>Audencia Nantes, School of Management, 8 route de la Jonelière, 44312 Nantes, France. Email: acharles@audencia.com.

<sup>†</sup>Corresponding author: LEMNA, University of Nantes, IEMN–IAE, Chemin de la Censive du Tertre, BP 52231, 44322 Nantes, France. Email: olivier.darne@univ-nantes.fr.

## 1 Introduction

Since the end of the 1990s oil prices have been steadily increasing, reflecting rising demand for crude oil, particularly from developing nations. Indeed, in the last decade the oil prices increased more than 700% (in average) and more than 80% (in average) between mid 2007 to mid 2008 due to rising demand, low spare capacity, weak dollar and geopolitical concerns (especially, tensions in Turkey, Nigeria and Iran). Oil prices have been very volatile, changing their trajectories and behavior with respect to the economic situation. Oil prices exhibit large upward or downward swings primarily caused by fluctuations in demand, extraction costs, and reserves (Pindyck, 1999). Supply and demand remain the main factors determining oil prices. More precisely, oil demands depend on oil consumption by developed and developing countries, and oil supplies depend on geopolitical events, such as troubles between Venezuela and the US or Turkey and Kurdish Iraq or Iran and Israel, among others, as well as oil tank levels and the Organization of the Petroleum Exporting Countries (OPEC) decisions on adjusting production levels. However, oil investor behavior can be also a factor in the recent price behavior, especially increasingly speculative behavior of a more diverse set of investors, including hedge funds, pension funds, and investment banks. All these factors question on the issue of whether or not the crude oil markets are predictable and therefore efficient.

In this paper we analyze the efficiency of the crude oil markets. The literature on market efficiency and stock market predictability is vast, as researchers have been discussing this theme in depth for the past decades (see Fama, 1970; 1991; Fama and French, 1988; Lo and MacKinlay, 1988; among others). A capital market is considered as efficient if stock prices at any time fully reflect all available and relevant information. Therefore, given only past price and return data, the current price is the best predictor of the future price, and the price change or return is expected to be zero. Stock prices exhibit no serial dependencies, meaning that there are no patterns to asset prices. This implies that future price movements are determined entirely by information not contained in the price series. This is the essence of the weak-form efficient market hypothesis [EMH], which implies a random walk. It is this random-walk implication of the weak-form EMH which is most commonly tested in the empirical literature<sup>1</sup>.

---

<sup>1</sup>Note that if the random walk hypothesis is based on the theory of efficiency, the EMH does not imply that prices follow a random walk. Therefore, if prices do not follow a random walk, this does not imply inefficiency of the market. See Lo and MacKinlay (2001) for a discussion on random walk hypothesis and efficiency market hypothesis.

Recently, Tabak and Cajueiro (2007), Alvarez-Ramirez et al. (2008) and Maslyuk and Smyth (2009) investigated the efficiency of crude oil markets from time-varying long-range dependence, Hurst exponent dynamics (from detrended fluctuation analysis) and unit root tests, respectively. They found that these markets are weak-form efficient. In this paper, we extend the examination of the weak-form of the EMH in the crude oil markets in two ways. First, this study is based on a more extensive sample. We study daily data for two crude oil markets, namely, the US West Texas Intermediate and the UK Brent, over the period June 1982 – July 2008. We also investigate the EMH over two sub-periods in order to analyze the effects of the important structural change due to policy changes that attempted to increase the efficiency of the North American energy industry in 1993<sup>2</sup>. Second, the weak form of the EMH is evaluated from an alternative method relative to the previous studies, namely the variance-ratio [VR] tests which are widely used in financial empirical studies (e.g., Lagoarde-Segot and Lucey, 2007; Kim and Shamsuddin, 2008)<sup>3</sup>. More precisely, we adopt individual non-parametric VR tests suggested by Wright (2000) as well as its multiple version proposed by Belaire-Franch and Contreras (2004). These VR tests are robust to heteroscedasticity and non-normality which are features displayed by the oil crude prices (Pindyck, 2004; Narayan and Narayan, 2007; Kang, Kang and Yoon, 2009), and powerful against fractionally integrated alternatives which are present in oil crude prices (Alvarez-Ramirez et al., 2002, 2008; Tabak and Cajueiro, 2007). We also apply the wild bootstrap VR tests suggested by Kim (2006) which are robust to heteroscedasticity and do not rely on asymptotic approximations.

The rest of this paper is organized as follows. Section 2 discusses the variance ratio tests. Section 3 summarizes the characteristics of the data on the Brent and WTI crude oil markets. Section 4 reports the empirical results. Section 5 concludes.

## 2 Variance ratio tests

Since the seminal work of Lo and MacKinlay (1988, 1989) and Poterba and Summers (1988), the standard variance ratio [VR] test or its improved modifications have been

---

<sup>2</sup>See Serletis and Andreadis (2004) and Serletis and Rangel-Ruiz (2004) for a discussion on these policy changes.

<sup>3</sup>Lo and MacKinlay (1989) examined the VR, Dickey-Fuller unit root and Box-Pierce serial correlation tests which are often employed to test the weak-form efficiency (Hoque, Kim and Pyun, 2007; Kim and Shamsuddin, 2008) and found that VR test was more powerful than the others under the heteroscedastic random walk.

widely used for testing market efficiency<sup>4</sup>.

The VR methodology consists of testing the random walk hypothesis [RWH] against stationary alternatives, by exploiting the fact that the variance of random walk increments is linear in all sampling intervals, i.e., the sample variance of  $k$ -period return (or  $k$ -period differences),  $y_t - y_{t-k}$ , of the time series  $y_t$ , is  $k$  times the sample variance of one-period return (or the first difference),  $y_t - y_{t-1}$ . The VR at lag  $k$  is then defined as the ratio between  $(1/k)$ th of the  $k$ -period return (or the  $k$ th difference) to the variance of the one-period return (or the first difference). Hence, for a random walk process, the variance computed at each individual lag interval  $k$  ( $k = 2, 3, \dots$ ) should be equal to unity.

In testing the null hypothesis of random walk, the VR test evaluates the hypothesis that a given time series or its first difference (or return),  $x_t = y_t - y_{t-1}$ , is a collection of independent and identically distributed observations (i.i.d.) or that it follows a martingale difference sequence. Following Wright (2000), the VR statistic be written as

$$VR(x; k) = \left\{ (Tk)^{-1} \sum_{t=k}^T (x_t + \dots + x_{t-k+1} - k\hat{\mu})^2 \right\} \div \left\{ T^{-1} \sum_{t=1}^T (x_t - \hat{\mu})^2 \right\} \quad (1)$$

where  $\hat{\mu} = T^{-1} \sum_{t=1}^T x_t$ . If the stock return follows a random walk, the expected value of  $VR(x; k)$  should be equal to unity for all horizons  $k$ . If this ratio is less than one at long horizons, then we have indications of negative serial correlation (mean-reversion) and ratios greater than one at long horizons implies positive serial correlation (mean-aversion).

Lo and MacKinlay (1988) proposed the asymptotic distribution of  $VR(x; k)$  by assuming that  $k$  is fixed when  $T \rightarrow \infty$ . They show that under the assumption of conditional heteroscedasticity, then under the null hypothesis that  $V(k) = 1$ , the test statistic  $M(x; k)$  is given by<sup>5</sup>

$$M(x; k) = \frac{VR(x; k) - 1}{\phi^*(k)^{1/2}} \quad (2)$$

---

<sup>4</sup>See Hoque, Kim and Pyun (2007) and Charles and Darné (2009) for a review.

<sup>5</sup>Lo and MacKinlay (1988) also propose a test statistic under the assumption of homoscedasticity. We focus only on VR statistic which is robust under heteroscedasticity since, as shown in Section 3, all the data display heteroscedasticity.

follows the standard normal distribution asymptotically, where

$$\begin{aligned}\phi^*(k) &= \sum_{j=1}^{k-1} \left[ \frac{2(k-j)}{k} \right]^2 \delta(j) \\ \delta(j) &= \left\{ \sum_{t=j+1}^T (x_t - \hat{\mu})^2 (x_{t-j} - \hat{\mu})^2 \right\} \div \left\{ \left[ \sum_{t=1}^T (x_t - \hat{\mu})^2 \right]^2 \right\}\end{aligned}$$

## 2.1 Wright (2000) tests

A well-known problem with the VR test is that the standard VR tests such as Lo and MacKinlay (1988) tests, which are based on asymptotic approximations, are biased (severe size distortions and low power) and right-skewed in finite samples, resulting in misleading statistical inference. Wright (2000) proposed a non-parametric alternative to conventional asymptotic VR tests using ranks. Wright's (2000) tests have two advantages over Lo-MacKinlay test when sample size is relatively small: (i) as the rank ( $R_1$  and  $R_2$ ) tests have exact sampling distribution, there is no need to resort to asymptotic distribution approximation, and (ii) the tests may be more powerful than the conventional VR tests against a wide range of models displaying serial correlation, including fractionally integrated alternatives. The tests based on ranks are exact under the i.i.d. assumption. Moreover, Wright (2000) showed that rank-based tests display low size distortion, under conditional heteroscedasticity<sup>6</sup>.

Given  $T$  observations of first differences of a variable,  $\{x_1, \dots, x_T\}$ , and let  $r(x)$  be the rank of  $x_t$  among  $(x_1, \dots, x_T)$ . Under the null hypothesis that  $x_t$  is generated from an i.i.d. sequence,  $r(x)$  is a random permutation of the numbers of  $1, \dots, T$  with equal probability. Wright (2000) suggested the  $R_1$  and  $R_2$  statistics, defined as

$$R_1(k) = \left( \frac{(Tk)^{-1} \sum_{t=k}^T (r_{1,t} + \dots + r_{1,t-k+1})^2}{T^{-1} \sum_{t=k}^T r_{1,t}^2} - 1 \right) \times \phi(k)^{-1/2} \quad (3)$$

$$R_2(k) = \left( \frac{(Tk)^{-1} \sum_{t=k}^T (r_{2,t} + \dots + r_{2,t-k+1})^2}{T^{-1} \sum_{t=k}^T r_{2,t}^2} - 1 \right) \times \phi(k)^{-1/2} \quad (4)$$

where the standardized ranks  $r_{1,t}$  and  $r_{2,t}$  are given by

$$\begin{aligned}r_{1,t} &= \frac{r(x_t) - \frac{T+1}{2}}{\sqrt{(T-1)(T+1)/12}} \\ r_{2,t} &= \Phi^{-1} \frac{r(x)}{T+1}\end{aligned}$$

---

<sup>6</sup>Wright (2000) also suggested sign-based tests.

where  $\phi(k) = 2(2k-1)(k-1)(3kT)^{-1}$ , and  $\Phi^{-1}$  is the inverse of the standard normal cumulative distribution function. The  $R_1$  and  $R_2$  statistics follow the same exact sampling distribution. The critical values of these tests can be obtained by simulating their exact distributions.

## 2.2 Belaire-Franch and Contreras (2004) tests

It is preferable to employ multiple VR tests rather than individual VR tests. Indeed, in practice, it is customary to examine the VR statistics for several holding periods ( $k$ ) to test the RWH. The null is rejected if it is rejected for some  $k$  value. As stressed by Chow and Denning (1993), this sequential procedure leads to size distortions. To overcome this problem, Chow and Denning (1993) provide a multiple VR test for the joint null hypothesis  $V(k_i) = 1$  for  $i = 1, \dots, m$  against the alternative that  $V(k_i) \neq 1$  for some holding period  $k_i$ . The (heteroscedasticity-robust) test statistic can be written as

$$MV(x; k_i) = \sqrt{T} \max_{1 \leq i \leq m} |M(x; k_i)| \quad (5)$$

where  $M(x; k_i)$  is defined in (2). Recently, Belaire-Franch and Contreras (2004) proposed a multiple rank VR tests by substituting Wright's rank-based tests in the definition of the Chow-Denning (1993) multiple test procedure<sup>7</sup>.

The statistics are defined as

$$\begin{aligned} CD_{(R_1)} &= \max_{1 \leq i \leq m} |R_1(k_i)| \\ CD_{(R_2)} &= \max_{1 \leq i \leq m} |R_2(k_i)| \end{aligned}$$

The rank-based procedures are exact under the i.i.d. assumption.

## 2.3 Kim (2006) test

Kim (2006) uses the wild bootstrap which is a resampling method that approximates the sampling distribution of the VR test statistic, and is applicable to data with unknown forms of conditional and unconditional heteroscedasticity (Mammen, 1993; Davidson and Flachaire, 2001; MacKinnon, 2002).

The wild bootstrap is applied to Chow-Denning,  $MV(x; k_i)$ , VR test. The wild bootstrap test based on  $MV(x; k_i)$  can be conducted in three stages as below

---

<sup>7</sup>Belaire-Franch and Contreras (2004) also suggested multiple sign-based VR tests based on Wright's sign-based tests but they showed that the rank-based tests are more powerful than the sign-based tests.

- (i) Form a bootstrap sample of  $T$  observations  $x_t^* = \eta_t x_t$  ( $t = 1, \dots, T$ ) where  $\eta_t$  is a random sequence with  $E(\eta) = 0$  and  $E(\eta^2) = 1$ .
- (ii) Calculate  $MV^* = MV(x^*; k_i)$ , the  $MV(x; k_i)$  statistic obtained from the bootstrap sample generated in stage (i).
- (iii) Repeat (i) and (ii) sufficiently many, say  $m$ , times to form a bootstrap distribution of the test statistic  $\{MV(j)\}_{j=1}^m$ .

The bootstrap distribution  $\{MV(j)\}_{j=1}^m$  is used to approximate the sampling distribution of the  $MV(x; k_i)$  statistic. The  $p$ -value of the test can be obtained as the proportion of  $\{MV(j)\}_{j=1}^m$  greater than the  $MV(x; k_i)$  statistic calculated from the original data. The wild bootstrap version of Lo-MacKinlay  $M(k)$  test can be implemented in a similar manner as a two-tailed test, where we obtain  $M^* = M(x^*; k)$  in stage (ii) and  $\{M(x^*; k_j; j)\}_{j=1}^m$  in stage (iii).

Conditionally on  $X_t$ ,  $X_t^*$  is a serially uncorrelated sequence with zero mean and variance  $X_t^2$ . As such,  $M^*$  and  $MV^*$  have the same asymptotic distributions as  $M_2(k)$  and  $MV_2(k_i)$ , respectively. Since  $X_t^*$  is a serially uncorrelated sequence, wild bootstrapping approximates the sampling distributions under the null hypothesis, which is a desirable property for a bootstrap test.

To implement the wild bootstrap test, a specific form of  $\eta_t$  should be chosen. Kim (2006) recommends using the standard normal distribution for  $\eta_t$  since he reports that other choices provided qualitatively similar small sample results<sup>8</sup>.

### 3 Data and summary statistics

The data of the study consists of the daily closing spot prices for two oil crude markets: the US West Texas Intermediate (WTI) and the UK Brent. The data comes from Thomson Financial Datastream and is given in US dollar per barrel. The data spans from 1 June, 1982 to 4 July, 2008, namely 6809 observations. Figures 1 and 2 provide a graphical representation of these series.

The WTI and Brent crude oils are traded on New York Mercantile Exchange (NYMEX) and InterContinental Exchange (ICE, London), respectively. The WTI and

---

<sup>8</sup>Kim (2006) finds that his simulation results are largely insensitive to different choices of  $\eta_t$ , including the two-point distribution of Mammen (1993), and the Rademacher distribution discussed in Davidson and Flachaire (2001).

Brent crude oils are light crude oils, with WTI lighter than Brent (the API<sup>9</sup> gravity is around 39.6 degrees for WTI against around 38.3) and sweet crude oils, with Brent slightly less sweet than WTI (Brent contains 0.37% of sulfur against 0.24% for WTI). Due to its very high quality, WTI crude oil is excellent for refining a larger portion of gasoline whereas Brent is ideal for making gasoline and middle distillates. The Brent crude oil is actually a combination of crude oil from 15 different oil fields in the Brent and Ninian systems located in the North Sea. Most WTI crude oil gets refined in the Midwest region of the US, more precisely at Cushing (Oklahoma), with some more refined within the Gulf Coast region. Brent crude oil is typically refined in Northwest Europe.

Table 1 presents summary statistics for the WTI and Brent crude oil returns calculated as the first differences in the logs of the spot prices. The two crude oil markets have approximately equal mean returns of about 0.02% per day, with the WTI returns actually marginally smaller than the Brent returns. All the returns are highly non-normal, i.e. showing evidence of significant negative skewness and excess kurtosis, as might be expected from daily returns. The Lagrange Multiplier test for the presence of the ARCH effect indicates clearly that all crude oil prices show strong conditional heteroscedasticity, which is a common feature of financial data.

## 4 Testing the efficient market hypothesis

We investigate the weak-form efficient market hypothesis by testing the random-walk hypothesis from VR tests. Table 2 reports the individual and multiple non-parametric VR test results. The test statistics are displayed for the Wright ( $R_1$  and  $R_2$ ) and Belaire-Contreras ( $CD(R_1)$  and  $CD(R_2)$ ) tests while the  $p$ -values are reported for the Kim ( $M^*$  and  $MV^*$ ) tests. The holding periods ( $k_i$ 's) considered are (2, 5, 10, 30). As advocated by Deo and Richardson (2003), we use relatively short holding periods when testing for the mean reversion using VR tests.

---

<sup>9</sup>The American Petroleum Institute (API) gravity is a measure of how heavy or light a petroleum liquid is compared to water. If its API gravity is greater than 10, it is lighter and floats on water; if less than 10, it is heavier and sinks. API gravity is used to compare the relative densities of petroleum liquids.

Figure 1: Daily Brent crude oil prices

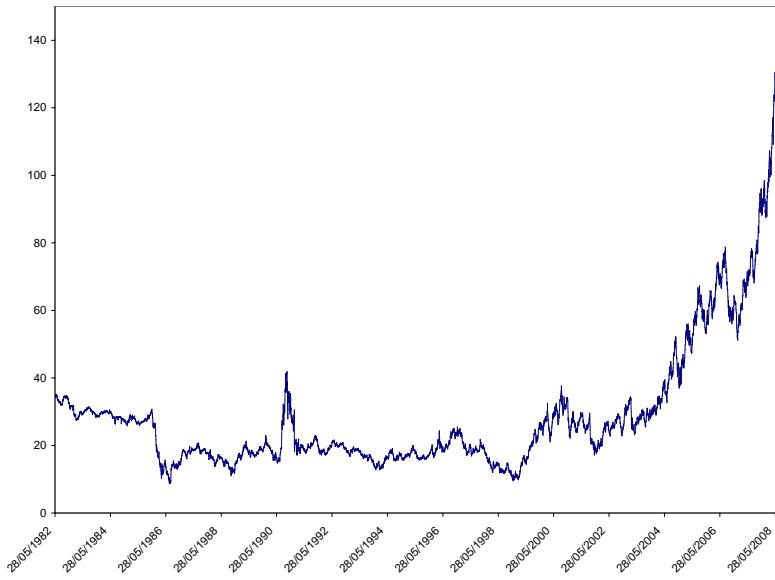
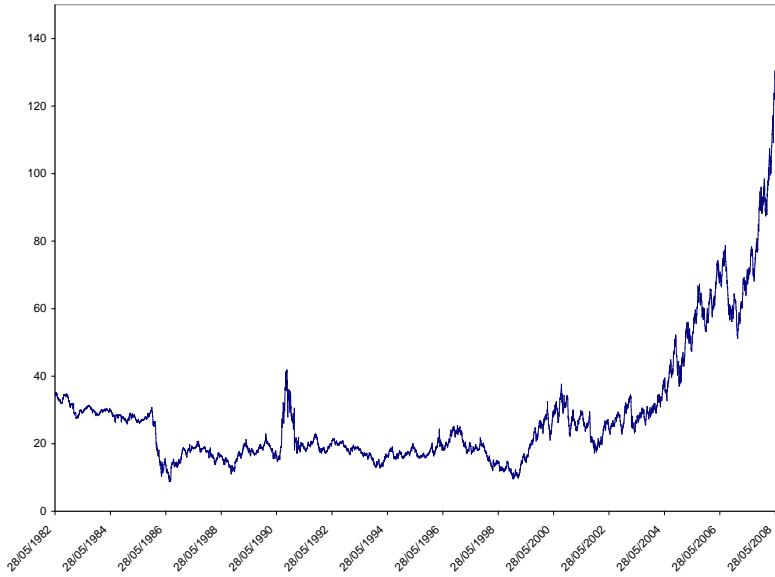


Figure 2: Daily WTI crude oil prices



Generally, the results are consistent with the RWH in the two crude oil prices, and therefore with weak-form efficiency of the two crude oil markets since the test statistics are not significant at the 5% level. Nevertheless, the multiple VR tests reject the RWH on the WTI market at the 10% level for the Belaire-Contreras tests and at the 5% level for the Chen-Deo test, implying some cautions on the EMH for this market.

In 1993 the North American Free Trade Agreement (NAFTA) signed by the United States, Canada and Mexico, has underpinned the process of deregulation and attempted to increase the efficiency of the North American energy industry (Serletis and Andreadis, 2004; Serletis and Rangel-Ruiz, 2004). It came into effect on January 1, 1994.

Therefore, we investigate whether the crude oil market has become more efficient since the early 1990s, namely the deregulated period, relative to the 1980s, namely the regulated period, by re-running the VR tests for the WTI and Brent crude oil prices for the following subperiods: 1 June, 1982 to 31 December, 1993, and 1 January, 1994 to 4 July, 2008.

Table 3 presents summary statistics for the two crude oil markets on the two sub-periods. For the Brent market, the returns on the 1982–1993 sub-period are higher (-0.021% per day) than those on the second sub-period (-0.031% per day). In the contrast, the returns of the WTI market on the two sub-periods have approximately equal mean returns of about 0.03% per day. Note that the risk (in terms of the standard deviation) associated with the 1994–2008 sub-period for the two crude oil markets is higher (2.30% per day) than that associated with the 1982–1993 sub-period (around 1.0% per day).

On the whole, all the statistics increase over the two sub-periods, especially for the excess kurtosis and the conditional heteroscedasticity, suggesting an impact of the deregulation on the crude oil prices. Furthermore, the skewness is negative and high, implying an increase in the extreme losses for the investors. The kurtosis is very high for the second sub-period, implying that the variance of the crude oil prices is principally due to infrequent but extreme deviations.

Tables 4 and 5 display the results of individual and multiple VR tests<sup>10</sup>. For the Brent crude oil market, the VR tests show the efficiency on the two sub-periods. On the contrary, for the WTI crude oil market, the RWH is rejected on the 1994–2008 sub-period while it is not rejected for the 1982–1993 sub-period, suggesting that the deregulation have not improved the efficiency on the WTI crude oil market in the sense of making returns less predictable. Thus, the NAFTA deregulation does not seem to have a positive impact on the efficiency of the WTI crude oil market, contrarily to what was expected. Furthermore, the Brent crude oil market seems to be more weak-form efficiency than the WTI crude oil market since the process of deregulation. This

---

<sup>10</sup>We also studied the effect of the Gulf War of 1990 on the VR tests. The outcomes are similar to those on the 1982–1993 period.

result can be explained by the fact that the WTI price can be affected by its storage tank capacity and its infrastructure logistics (Horsnell and Mabro, 1993). For example, high levels of stocks at Cushing (e.g., due to unsatisfactory supply-demand conditions) tend to depress the WTI price relative to the Brent price because the WTI crude oil market suffers from pipeline logistic, more precisely infrastructure bottlenecks. These results are in contradiction with those of Tabak and Cajueiro (2007), finding that the WTI crude oil prices seems to be more weak-form efficient than Brent prices.

## 5 Conclusion

This paper has explored weak efficiency in the Brent and WTI crude oil markets and the relationship between efficiency and changes in the regulation, using new variance ratio tests. Theses tests, which are robust to heteroscedasticity and non-normality and powerful against fractionally integrated alternatives which are present in oil crude prices, are the non-parametric VR tests developed by Wright (2000) and Belaire-Franch and Contreras (2004) as well as the wild-bootstrap VR tests suggested by Kim (2006).

The results suggest that the Brent crude oil prices is weak-form efficiency while the WTI crude oil market seems to be inefficiency on the 1994–2008 sub-period, suggesting that the deregulation has not improved the efficiency on the WTI crude oil market in the sense of making returns less predictable. The storage tank capacity and infrastructure logistics of the WTI crude oil market can explain the fact that the Brent crude oil market is more efficient than the WTI crude oil market.

## References

- [1] Alvarez-Ramirez, J., Alvarez, J., Rodriguez E., 2008. Short-term predictability of crude oil markets: A detrended fluctuation analysis approach. *Energy Economics*, 30, 2645–2656.
- [2] Alvarez-Ramirez, J., Cisneros, M., Ibarra-Valdez, C., Soriano, A., 2002. Multifractal Hurst analysis of crude oil prices. *Physica A*, 313, 651–670.
- [3] Belaire-Franch, J., Contreras, D., 2004. Ranks and signs-based multiple variance ratio tests. Working paper, Department of Economic Analysis, University of Valencia.
- [4] Charles, A., Darné, O., 2009. Variance ratio tests of random walk: An overview. *Journal of Economic Surveys*, forthcoming.
- [5] Chow, K.V., Denning, K.C., 1993. A simple multiple variance ratio test. *Journal of Econometrics*, 58, 385–401.
- [6] Davidson, R., Flachaire, E., 2001. The wild bootstrap, tamed at last. Working paper No 99, GREQAM, University of Marseille.
- [7] Deo, R.S., Richardson, M., 2003. On the asymptotic power of the variance ratio test. *Econometric Theory*, 19, 231–239.
- [8] Fama, E.F., 1970. Efficient capital markets: A review of theory and empirical work. *Journal of Finance*, 25, 383–417.
- [9] Fama, E.F., 1991. Efficient capital markets: II. *Journal of Finance*, 46, 1575–1617.
- [10] Fama, E.F., French, K.R., 1988. Permanent and temporary components of stock prices. *Journal of Political Economy*, 96, 246–273.
- [11] Hoque, H.A.A.B., Kim, J.H., Pyun, C.S., 2007. A comparison of variance ratio tests of random walk: A case of Asian emerging stock markets. *International Review of Economics and Finance*, 16, 488–502.
- [12] Horsnell, P., Mabro, R., 1993. *Oils Markets and Prices: The Brent Market and the Formation of World Oil Prices*. Oxford University Press, Oxford.
- [13] Kang, S.H., Kang, S-M., Yoon, S-M., 2009. Forecasting volatility of crude oil markets. *Energy Economics*, 31, 119–125.

- [14] Kim, J.H., 2006. Wild bootstrapping variance ratio tests. *Economics Letters*, 92, 38–43.
- [15] Kim, J.H., Shamsuddin, A., 2008. Are Asian stock markets efficient? Evidence from new multiple variance ratio tests. *Journal of Empirical Finance*, 15, 518–532.
- [16] Lagoarde-Segot, T., Lucey, B.M., 2007. Efficiency in emerging markets: Evidence from the MENA region. *International Financial Markets, Institutions and Money*, 18, 94–105.
- [17] Lo, A.W., MacKinlay, A.C., 1988. Stock market prices do not follow random walk: Evidence from a simple specification test. *Review of Financial Studies*, 1, 41–66.
- [18] Lo, A.W., MacKinlay, A.C., 1989. The size and power of the variance ratio test in finite samples: A Monte Carlo investigation. *Journal of Econometrics*, 40, 203–238.
- [19] Lo, A.W., MacKinlay, A.C., 2001. *A Non-Random Walk Down Wall Street*. Princeton University Press, Princeton.
- [20] Mammen, E., 1993. Bootstrap and wild bootstrap for high dimensional linear models. *Annals of Statistics*, 21, 255–285.
- [21] MacKinnon, J.G., 2002. Bootstrap inference in econometrics. *Canadian Journal of Economics*, 35, 615–645.
- [22] Maslyuk, S., Smyth, R., 2009. Unit root properties of crude oil spot and futures prices. *Energy Policy*, 31, 109–118.
- [23] Narayan, P.K., Narayan, S., 2007. Modelling oil price volatility. *Energy Policy*, 35, 6549–6553.
- [24] Pindyck, R.S., 1999. The long-run evolution of energy prices. *Energy Journal*, 20, 1–27.
- [25] Pindyck, R.S., 2004. Volatility in natural gas and oil market. *Journal of Energy and Development*, 30, 1–19.
- [26] Serletis, A., Andreadis, I., 2004. Random fractal structures in North American energy markets. *Energy Economics*, 26, 389–399.

- [27] Serletis, A., Rangel-Ruiz, R., 2004. Testing for common features in North American energy markets. *Energy Economics*, 26, 401–414.
- [28] Tabak, B.M., 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.
- [29] Wright, J.H., 2000. Alternative variance-ratio tests using ranks and signs. *Journal of Business and Economic Statistics*, 18, 1–9.

Table 1: Statistical analysis of returns

Market	Obs.	Mean (% daily)	SD (% daily)	Skewness	Kurtosis	LB(10)	ARCH(10)
Brent	6810	0.0209	2.20	-1.031*	30.86*	41.30*	158.16*
WTI	6810	0.0203	2.30	-1.040*	25.63*	40.02*	206.31*

Notes: The skewness and kurtosis statistics are standard-normally distributed under the null of normality distributed returns. ARCH(10) indicates the Lagrange multiplier test for conditional heteroscedasticity with 10 lags. LB(10) corresponds to the Ljung-Box test for serial autocorrelation with 10 lags. \* and \*\* means significant at 1% and 5% level, respectively.

Table 2: Results of individual and multiple VR tests – 1982–2008.

		$k$			
		2	5	10	30
Brent	$R_1$	0.903	-0.233	0.254	1.724
	$R_2$	0.836	-0.354	-0.356	1.111
	$M^*$	0.494	0.578	0.255	0.776
	$CD(R_1)$	1.270			
	$CD(R_2)$	0.852			
	$MV^*$	0.529			
WTI	$R_1$	-1.386	-1.749	-1.190	0.497
	$R_2$	-1.274	-2.064*	-2.165*	-0.499
	$M^*$	0.717	0.071**	0.012*	0.138
	$CD(R_1)$	1.974**			
	$CD(R_2)$	2.236**			
	$MV^*$	0.037*			

\* and \*\* Significant at the 5% and 10% level, respectively. We report the VR statistic for each test.

Table 3: Statistical analysis of returns

Market	Obs.	Mean (% daily)	SD (% daily)	Skewness	Kurtosis	LB(10)	ARCH(10)
<b>Brent</b>							
1982–1993	937	-0.021	0.90	-1.686*	21.41*	40.49*	69.07*
1994–2008	3025	-0.031	2.30	-1.913*	55.39*	18.05*	142.99*
<b>WTI</b>							
1982–1993	937	-0.033	1.00	-1.519*	38.71*	35.93*	95.81*
1994–2008	3025	-0.030	2.30	-1.882*	45.18*	21.71*	112.52*

Notes: The skewness and kurtosis statistics are standard-normally distributed under the null of normality distributed returns. ARCH(10) indicates the Lagrange multiplier test for conditional heteroscedasticity with 10 lags. LB(10) corresponds to the Ljung-Box test for serial autocorrelation with 10 lags. \* and \*\* means significant at 1% and 5% level, respectively.

Table 4: Results of individual and multiple VR tests – Brent.

		<i>k</i>			
		2	5	10	30
1982–1993	$R_1$	1.831*	0.680	0.908	2.609*
	$R_2$	1.920*	0.617	0.394	2.065*
	$M^*$	0.274	0.912	0.402	0.896
	$CD(R_1)$	1.876			
	$CD(R_2)$	1.962			
	$MV^*$	0.503			
1994–2008	$R_1$	-0.350	-1.043	-0.688	-0.226
	$R_2$	-0.564	-1.190	-0.988	-0.419
	$M^*$	0.470	0.238	0.265	0.595
	$CD(R_1)$	1.187			
	$CD(R_2)$	1.257			
	$MV^*$	0.517			

\* and \*\* Significant at the 5% and 10% level, respectively. We report the VR statistic for each test.

Table 5: Results of individual and multiple VR tests – WTI.

		<i>k</i>			
		2	5	10	30
1982–1993	$R_1$	0.862	0.103	0.561	2.109*
	$R_2$	0.601	-0.400	-0.445	1.376
	$M^*$	0.940	0.171	0.054*	0.509
	$CD(R_1)$	0.959			
	$CD(R_2)$	0.686			
	$MV^*$	0.152			
1994–2008	$R_1$	-2.187*	-2.346*	-2.007*	-1.253
	$R_2$	-1.811	-2.242*	-2.303*	-1.828*
	$M^*$	0.362	0.118	0.030*	0.051*
	$CD(R_1)$	2.569*			
	$CD(R_2)$	2.527*			
	$MV^*$	0.113			

\* and \*\* Significant at the 5% and 10% level, respectively. We report the VR statistic for each test.