Buberkoku, Onder

Article

Examining energy futures market efficiency under multiple regime shifts

International Journal of Energy Economics and Policy

Reference: Buberkoku, Onder (2017). Examining energy futures market efficiency under multiple regime shifts. In: International Journal of Energy Economics and Policy 7 (6), S. 61 - 71.

International Journal of Energy Economics and Policy

ISSN: 2146-4553

available at http: www.econjournals.com

International Journal of Energy Economics and Policy, 2017, 7(6), 61-71.

Examining Energy Futures Market Efficiency under Multiple Regime Shifts

Onder Buberkoku*

Department of Finance, Faculty of Business Administration, Yuzuncu Yil University, Van, Turkey. *Email: onderbuber@gmail.com

ABSTRACT

This study examines the West Texas Intermediate crude oil (WTI), Europe Brent crude oil (Brent), heating oil no. 2, and Henry Hub natural gas (NG) futures markets' efficiency following Fama's (1970) weak-form efficiency hypothesis, using spot and futures prices at 1, 2, 3, and 4 months maturity based on the tests with unknown multiple regime shifts. The results show that it is important to consider the multiple regime shifts when determining whether energy futures markets are efficient. We find that WTI and Brent futures markets are not efficient, whereas NG and heating oil futures markets are efficient. Additionally, the findings also shed light on discussions about the stationary properties of energy commodities and whether spot and futures prices are cointegrated. In particular, this study presents new evidence based on the unit root and cointegration tests with multiple structural breaks.

Keywords: Energy Commodity, Futures Market Efficiency, Multiple Structural Breaks
JEL Classifications: G14, G15, Q40

1. INTRODUCTION

Changes in the energy commodity prices can lead to a significant impact on the global economy (Sadorsky, 2006). Therefore, the volatility in the energy market is an important issue for policy makers, producers, as well as risk managers. Particularly in the last decade, such developments as increasing speculative trading and the US dollar fluctuations have caused a significant increase in the volatility of energy commodities (Fan and Xu, 2011). Therefore, accurately predicting energy commodity prices and hedging their market risk have become crucial. Moreover, futures market is one of the tools that can be used for forecasting future spot prices and managing the market risk of energy commodities. In other words, futures markets have two main functions: Risk management and price discovery. However, the ability of futures markets to accurately fulfil these functions depends on whether the futures markets are efficient. Based on the efficient market hypothesis developed by Fama (1970), for a futures market to be efficient, it should fully reflect all available information about the underlying assets. In other words, the futures prices should be unbiased predictors of the future spot prices.

There is much literature examining futures market efficiency for energy commodities. The results, however, are mixed. For example, Lee and Zeng (2011) investigate the West Texas Intermediate (WTI) futures market efficiency under different maturities of futures contracts using quantile cointegration regression. They find that maturities of futures contracts affect the cointegration relationship between spot and futures oil prices, and only short maturities futures contracts are consistent with the efficient market hypothesis. Switzer and El-Khoury (2007) examine the efficiency of the NYMEX light sweet crude oil futures markets employing the Johansen (1988) cointegration test, and report that their results support market efficiency even during the episodes of extreme conditional volatility. Arouri et al. (2013) analyse the efficiency of nine energy and precious metal futures markets applying both linear and non-linear econometric techniques to test both long- and short-run efficiencies. They indicate that although futures prices are cointegrated with spot prices, they are not unbiased predictors of future spot prices. Shambora and Rossiter (2007) test the efficiency of crude oil futures contracts based on the artificial neural network model and they document that the crude oil futures market is not efficient because it presents profitable trading opportunities. Moosa and Al-Loughani (1994) perform several tests on market efficiency and unbiasedness hypotheses for WTI futures market. They indicate that futures prices are neither unbiased nor efficient forecasters of the spot prices.

Further, Peroni and Mcnown (1998) apply two informative tests to WTI, heating oil no.2, and unleaded gasoline futures markets, and reveal that the results are largely supportive of the efficiency hypothesis in three energy futures markets. Kawamoto and Hamori (2011) examine the market efficiency and unbiasedness among WTI futures with different maturities. They report that WTI futures market is efficient within an 8-month maturity, and efficient and unbiased within a 2-month maturity. Lean et al. (2010) test the WTI futures market efficiency using both mean-variance and stochastic dominance approaches and find that WTI futures market is efficient. Zhang and Wang (2013) explore the price discovery and risk transfer functions in crude oil and gasoline futures markets by using the model introduced by Garbade and Silber (1983). They reveal that while crude oil futures markets perform well in both the price discovery and risk transfer functions, gasoline futures market perform well in only the price discovery function. Gebre-Mariam (2011) employs the causality and cointegration tests to analyse the efficiency of natural gas (NG) market and reveals that the market efficiency holds only for contracts with about 1 month to maturity. Abosedra and Baghestani (2004) evaluate the WTI futures market efficiency using the 1-, 3-, 6-, 9-, and 12-month-ahead futures prices. They reveal that all the relevant crude oil futures prices are unbiased predictors of future spot prices. Beck (1994) applies traditional cointegration techniques to test the futures market efficiency for five commodity markets. He reports that all five markets are sometimes inefficient, but no market is always inefficient. Crowder and Hamid (1993) evaluate crude oil futures market efficiency based on cointegration analysis and obtain results that support the simple efficiency hypothesis.

Despite these and similar studies, a limitation of the relevant literature, as Maslyuk and Smyth (2009) and Chen et al. (2014) point out, is that few studies have so far considered the impact of structural breaks on energy commodities futures market efficiency. However, as widely reported in the literature, allowing for potential structural changes in economic process is an important issue (Hatemi-J, 2008). Financial crises, technological advances, policy changes, economic agents' behaviour, and external shocks may cause structural breaks. In this regard, when we consider the last 15 years of energy commodities, developments such as the 2001 dot-com bubble crisis, 2003 Iraq War, 2007–2008 subprime mortgage crisis, and 2011 Arab Spring may have caused structural breaks. Besides, as Fan and Xu (2011), among others, point out, since 1999–2000, the energy commodity market has undergone significant changes, and the increasing demand of emerging markets and growing financialisation and liberalisation of commodity markets are among the main factors leading to these changes. Furthermore, examining oil price dynamics, Askari and Krichene (2008) document that, even in the short run, there are large price changes in the oil market. Therefore, all these discussions indicate that in analysing energy commodities, it is important to consider potential structural breaks. Indeed, in recent literature, there are some studies allowing multiple breaks for energy commodities (Lee and Lee, 2009; Noguera, 2013).[1]

1 Moreover, energy commodity series' plots also imply that series may have multiple breaks, especially in their level and slope of time trend. However, since the series' plots are commonly shown in the literature, they are not presented in this paper, but are avaliable upon request.

In this study, we aim to examine the WTI, Brent, heating oil (HT hereafter), and NG futures markets' efficiency following Fama's (1970) weak-form efficiency hypothesis under the possible multiple structural breaks. In this context, first, the Gregory and Hansen (1996) (GH hereafter) cointegration test allowing one unknown regime shift and the Hatemi-J (2008) (HJ hereafter) cointegration test allowing two unknown regime shifts are used. Then, a new cointegration test developed by Maki (2012) allowing unknown breaks up to five is employed. The reason for following such a methodological procedure is because cointegration is a necessary condition for market efficiency. However, as pointed out by Maki (2012), we generally do not have a priori information about the true number of breaks. Therefore, if the true number of breaks is two, then the GH test is misspecified, which will lead to a poor performance. Similarly, if the true number is one, the HJ test will suffer from the same problem. Additionally, if the true number is more than two, both tests will have a poor performance. Thus, it also makes sense to use the Maki (2012) test considering an unknown number of breaks. Further, Maki (2012) also shows that this newly developed test performs better than the GH and HJ tests when cointegration relationship has more than three breaks or persistent Markov switching.

The contributions of the study to the literature are as follows. First, as mentioned previously, although structural breaks are one of the characteristics of energy commodities (Lee and Lee, 2009), the existing literature has focussed less on this issue thus far when examining the futures market efficiency based on Fama's (1970) hypothesis. Therefore, this study fills this gap by employing tests allowing unknown multiple breaks. Second, studies generally analyse a limited number of energy commodities, namely WTI or Brent, and use only nearby futures contracts or one maturity level of futures contracts. However, in this study, in addition to WTI and Brent, we also examine HT and NG markets. Additionally, since only nearby or one maturity level of futures contract may not be sufficient to represent the whole futures market (Tang et al., 2013) and different levels of maturity are also important in terms of hedging and efficiency (Naraya et al., 2010), we use futures contracts at 1, 2, 3, and 4 months to maturity for each energy commodity. Third, the impact of multiple structural breaks on parameter estimates is also considered, which is generally ignored even in studies based on structural break analysis. Fourth, indirectly, this study also sheds light on discussions about the stationary properties of energy commodities and whether the spot and futures prices are cointegrated (Narayan and Liu, 2011; Ozdemir et al., 2013; Wang and Wu, 2013). In particular, this study presents new findings based on the unit root and cointegration tests by allowing five endogenous multiple structural breaks. To the best of our knowledge, these tests have not been applied to these series before.

The remainder of this paper is organised as follows: Section 2 presents the method and market efficiency hypothesis. Section 3 presents the empirical findings, while Section 4 concludes.

2. METHODOLOGY

2.1. Data

In this study, WTI, Brent, HO, and NG futures markets' efficiency are examined using weekly data for spot and futures prices at 1, 2, 3, and 4 months to maturity, covering the period from January 1, 1999 to November 29, 2013. All the data are extracted from the Energy Information Administration except for Brent futures data, which is from the Intercontinental Exchange.

2.2. Market Efficiency Hypothesis

Fama's (1970) weak-form efficiency hypothesis is tested based on the following model:

$$LS_t = \alpha_0+\beta_0 LF_{i,t}+\varepsilon_t,\ i = 1, 2, 3, 4 \text{ months to maturity} \quad (1)$$

Where LS and LF are the logarithmic spot and futures prices at time t, α_0 and β_0 are the model parameters, and ε is the error term. Based on the efficiency hypothesis, for futures prices to be an unbiased predictor of future spot prices, $\alpha_0 = 0$ and $\beta_0 = 1$ restrictions in Equation (1) should not be rejected jointly. If these restrictions are rejected, it means that either futures market is inefficient or that investors are not risk neutral, implying that they demand a constant or time-varying risk premium (Arouri et al., 2013). In this study, we also test $\alpha_0 = 0$ and $\beta_0 = 1$ restrictions separately, as ensuring $\beta_0 = 1$ restriction is crucial in terms of market efficiency. This is because $\alpha_0 = 0$ restriction may not hold if there exists a constant or time-varying risk premium or transportation costs even when futures markets are efficient (Chin et al., 2005; Kawamoto and Hamori, 2011; McKenzi and Holt, 2002; Wang and Ke, 2005). Therefore, as pointed out by Kawamoto and Hamori (2011), among others, $\beta_0 = 1$ restriction represents the null hypothesis of market efficiency, whereas $\alpha_0 = 0$ and $\beta_0 = 1$ joint restriction represents the null hypothesis of market efficiency and unbiasedness.

2.3. Structural Break and Unit Root Tests

We use the double maximum tests (UDmax and WDmax tests) proposed by Bai and Perron (1998; 2003) to detect whether spot and futures series have structural breaks. The UDmax and WDmax tests examine the null hypothesis of no structural breaks against the alternative hypothesis of an unknown number of breaks. Then, we employ the augmented Dickey–Fuller (1979) (ADF hereafter) unit root test to determine the integration order of the spot and futures series. Additionally, because standard unit root tests are biased towards the rejection of null hypothesis of unit root (Perron, 1989) in the presence of structural breaks, Zivot and Andrews' (1992) (ZA hereafter) endogenous structural break test is also performed. ZA propose three different models: Models A, B, and C, which allow a break in level, a break in slope, and a break in both level and slope, respectively. In our study, all three models are applied. However, one of the drawbacks of the ZA test is that it considers only one break, and as mentioned before, energy commodities may have multiple breaks. Therefore, Carrion-I-Silvestre et al.'s (2009) (CS hereafter) unit root test, which allows up to five breaks in the level and slope of time trend, is also employed. This test has five different test statistics, namely $MZ_\alpha^{GLS}(\lambda)$, $MSB^{GLS}(\lambda)$, $MZ_t^{GLS}(\lambda)$, $MP_T^{GLS}(\lambda)$, and $P_T^{GLS}(\lambda)$ tests, which are the so-called M-class of tests analysed by Ng and Perron (2003). Each test statistics has the null hypothesis of unit root. For this null hypothesis to be rejected, the estimated test statistics should be smaller than its critical values.

2.4. Cointegration Tests with Structural Breaks

Standard cointegration tests assume that the cointegration vector does not change over time. However, as pointed out by Lee and Lee (2009), energy commodity price series are usually affected by multiple breaks. Therefore, it is more appropriate to use cointegration tests that allow structural breaks. For this, GH proposes a test allowing cointegrating relationship to change. However, the GH test allows only one endogenous break. Therefore, HJ (2008) extended the GH test to account for two endogenous structural breaks. Both tests have three test statistics, namely the ADF*, Z_t^* and Z_α^* tests, to test the null hypothesis of no cointegration. Additionally, both tests consider three different structural change models: Level shift model (C), level shift with trend model (C/T), and regime shift model (C/S). To be consistent with the aim of the study, we apply the regime shift model for both tests. Based on Equation (1), this model can be defined for the GH and HJ tests as follows:

$$LS_t = \alpha_0+\alpha_1 D_{1t}+\beta_0 LS_t+\beta_1 D_{1t} LS_t+\varepsilon_t \quad (2)$$

$$LS_t = \alpha_0+\alpha_1 D_{1t}+\alpha_2 D_{2t}+\beta_0 LS_t+\beta_1 D_{1t} LS_t+\beta_2 D_{2t} LS_t+\varepsilon_t \quad (3)$$

Where α_0 and β_0 are the intercept and slope coefficients before the break, α_1 and α_2 are the changes in the intercept at the time of first and second breaks, and β_1 and β_2 are the changes in the slope at the time of first and second breaks. D_{1t} and D_{2t} are the dummy variables and defined as follows:

$$D_{1t} = \begin{Bmatrix} 0, t \le [\eta\tau 1] \\ 1, t > [\eta\tau 1] \end{Bmatrix};\ \text{and}\ D_{2t} = \begin{Bmatrix} 0, t \le [\eta\tau 2] \\ 1, t > [\eta\tau 2] \end{Bmatrix},$$

Where the unknown parameter τϵ (0,1) denotes the time of the break and [.] refers to the integer part.

However, a limitation of the HJ test is that it allows only two unknown breaks, so we also perform the Maki (2012) cointegration test considering up to five endogenous breaks[2]. This test, based on tests for structural breaks introduced by Bai and Perron (1998; 2003) and the unit root test with structural breaks developed by Kapetanious (2005), assumes that the unspecified number of breaks may be smaller than or equal to the maximum number of breaks set a priori. Besides, this test is considerably less computationally intensive than methods widely used in the literature (Maki, 2012).

The Maki (2012) test allows four different structural change models: Level shift model (C), level shift with trend model (C/T), regime shift model (C/S), and trend and regime shifts model (C/S/T), which allows changes in both level, trend and regressors. Based on Equation (1), the regime shift model is given by;

$$LS_t = \alpha_0 + \sum_{i=1}^{k} \alpha_i D_{it} + \beta_0 LF_t + \sum_{i=1}^{k} \beta_i D_{it} LF_t + \varepsilon_t \quad (4)$$

Where k is the maximum number of breaks, and D_{it} is the dummy variable defined as:

2 Recently, Çağli and Mandaci (2013) also use the Maki (2012) cointegration test while examining the long-run relationship between spot and futures prices under multiple regime shifts.

$$D_{it} = \begin{cases} 0 \text{ if } t \le T_{Bi}, i = 1, 2\&.5 \\ 1 \text{ if } t > T_{Bi}, i = 1, 2, \&..5 \end{cases}$$

Where, T_{Bi} is the time period of the break.

Similar to the GH and HJ tests, the Maki (2012) test is also a residual-based cointegration test with null hypothesis of no cointegration and alternative hypothesis of cointegration with *i* breaks ($i \le k$).

3. RESULTS

Tables 1 and 2 present the structural break and unit root test results. Both the UDmax and WD max test statistics reject the null hypothesis of no structural breaks in each case. This implies that there is at least one structural break in each of the series. Additionally, the ADF unit root test results show that spot and four different futures prices of all energy commodities have a unit in their level form, whereas the first differences of series are found to be stationary. However, one reason why the ADF unit root test is unable to reject the null hypothesis of unit root may be the presence of structural breaks. Therefore, we also employ the ZA and CS unit root tests allowing one and five structural breaks. In all cases, the CA unit root test finds five breaks in the level and slope of time trend in each series. Moreover, both the ZA and CS unit root tests show that all the series have unit roots in their level form at the 5% significance level. All these results indicate that all series are integrated of order one, I (1), and thus appropriate for cointegration analysis. They also shed light on discussions about the stationary properties of energy commodities, and show that the relevant energy commodity series are not stationary even if five structural breaks are allowed in the level and slope of time trend, indicating that shocks to these energy commodities will have a persistent effect on them. Finally, these findings are also consistent with the recent findings by Ozdemir et al. (2013) who adds to the relevant literature by allowing three breaks in the univariate time series models.

Having established that series are integrated of order one, the next step is to employ the cointegration tests. However, first, we check whether the model presented in Equation (1) has a regime shift. For this, we use the cumulative sum (CUSUM) and CUSUM of squares test statistics. The results show that in all cases, the models have regime shifts[3]. Then, we apply the GH and HJ cointegration tests; Table 3 shows the results. The results reveal that all three ADF*, Z_t^* and Z_α^* tests statistics of both the GH and HJ tests reject the null hypothesis of no cointegration at the 5% or a better significance level in all cases. This suggests that spot and four different futures prices of all energy commodities have cointegration relationship with regime shifts, and the maturity of futures contracts do not affect the cointegration relationship between the spot and futures prices.

Table 4 presents the Maki (2012) test results[4]. First, in all cases, the Maki (2012) test finds five regime shifts. However, the results

3 For simplicity, the results are not presented here, but are avaliable upon request.

4 We recognise a small error in Maki's (2012) original paper, namely that the order of models showing the critical values in Table 1 titled 'Critical values of cointegration tests with multiple breaks' (p. 2013) is wrong. The right order is that model 0, 1, 2, and 3 in Table 1 should correspond to model (C), model (C/T), model (C/S), and model (C/S/T), respectively. Additionally, to be sure, we sent an e-mail to Mr. Maki and he also verified this issue. Therefore, in this study, we use the critical values according to this order, which means that critical values for regime shift models at the 5% significance level is -6.357.

Table 1: Structural break and the ADF and ZA unit root tests results

Variables	UDmax	WDmax	ADF		ZA (level)		
			Level	First difference	Model A	Model B	Model C
WTI							
LS	74.866*	164.28*	−3.388 (3)	−14.106 (2)*	−4.918 (3)	−3.798 (3)	−4.834 (3)
LF1	75.348*	165.34*	−3.054 (1)	−23.748 (0)*	−4.566 (1)	−3.445 (1)	−4.466 (1)
LF2	79.643*	174.77*	−2.925 (1)	−23.402 (0)*	−4.403 (1)	−3.387 (1)	−4.326 (1)
LF3	85.110*	186.76*	−2.789 (1)	−23.353 (0)*	−4.269 (1)	−3.305 (1)	−4.199 (1)
LF4	89.425*	196.23*	−2.648 (1)	−23.304 (0)*	−4.139 (1)	−3.210 (1)	−4.074 (1)
Brent							
LS	209.71*	406.18*	−3.361 (0)	−27.426* (0)	−4.621 (0)	−3.516 (0)	−4.464 (0)
LF1	261.79*	574.46*	−3.260 (0)	−29.046* (0)	−4.532 (0)	−3.436 (0)	−4.349 (0)
LF2	280.38*	615.26*	−3.131 (0)	−29.304* (0)	−4.471 (0)	−3.341 (0)	−4.261 (0)
LF3	295.25*	647.88*	−2.985 (0)	−29.414* (0)	−4.380 (0)	−3.240 (0)	−4.214 (0)
LF4	302.49*	663.78*	−2.843 (0)	−29.438* (0)	−4.282 (0)	−3.143 (0)	−4.172 (0)
HT							
LS	64.99*	123.06*	−3.083 (2)	−20.372* (1)	−4.366 (2)	−3.271 (2)	−4.257 (2)
LF1	68.79*	138.92*	−2.814 (2)	−20.042* (1)	−4.243 (2)	−2.997 (2)	−4.079 (2)
LF2	67.64*	148.42*	−2.861 (1)	−19.341* (1)	−4.363 (1)	−3.142 (1)	−4.248 (1)
LF3	81.24*	178.27*	−2.779 (1)	−19.143* (1)	−4.320 (1)	−3.093 (1)	−4.196 (1)
LF4	151.08*	331.52*	−2.665 (1)	−23.314* (0)	−4.227 (1)	−3.003 (1)	−4.086 (1)
NG							
LS	12.05*	16.775*	−2.982 (1)	−23.978* (0)	−4.638 (1)	−4.318 (1)	−4.888 (1)
LF1	9.023*	14.581*	−2.681 (1)	−23.312* (0)	−4.469 (1)	−3.873 (1)	−4.541 (1)
LF2	8.197**	14.09*	−2.552 (1)	−22.775* (0)	−4.412 (1)	−3.799 (1)	−4.492 (1)
LF3	8.504**	14.623*	−2.475 (1)	−22.607* (0)	−4.410 (1)	−3.849 (1)	−4.576 (1)
LF4	8.805**	15.139*	−2.226 (1)	−23.387* (0)	−4.368 (1)	−3.542 (1)	−4.402 (1)

The figures in parentheses are the lag lengths. ADF unit root test is estimated with constant and trend. LF1, LF2, LF3, and LF4 are the futures prices at 1, 2, 3, and 4 months maturity, respectively. 15% trimming region is used for double maximum tests.* and ** denote significance at the 1% and 5% levels, respectively. WTI: West Texas Intermediate, ZA: Zivot and Andrews, ADF: Augmented Dickey–Fuller, NG: Natural gas

Table 2: Carrion-I-Silvestre *et al.* (2009) multiple structural break unit root test results

Variables	P_T^{GLS} (λ)	MP_T^{GLS} (λ)	MZ_α^{GLS} (λ)	MSB^{GLS} (λ)	MZ_t^{GLS} (λ)	m
WTI						
LNS	29.704	27.405	−15.212	0.1808	−2.7516	5
	8.9073	*8.9073*	*−46.5233*	*0.1032*	*−4.7995*	
LNF1	22.331	20.766	−20.193	0.1572	−3.1760	5
	8.9877	*8.9877*	*−46.4470*	*0.1035*	*−4.7902*	
LNF2	25.172	21.744	−19.3417	0.1608	−3.1096	5
	9.0353	*9.0353*	*−46.4379*	*0.1036*	*−4.7869*	
LNF3	26.619	22.973	−19.395	0.1605	−3.1140	5
	9.4540	*9.4540*	*−47.0014*	*0.1035*	−4.8006	
LNF4	25.244	23.0174	−19.3602	0.1607	−3.1107	5
	9.4149	*9.4149*	*−47.1842*	*0.1032*	*−4.8112*	
Brent						
LNS	14.481	13.085	−32.803	0.1234	−4.0488	5
	9.1748	*9.1748*	*−46.729*	*0.1034*	*−4.7996*	
LNF1	17.651	14.361	−28.938	0.1314	−3.8032	5
	8.9702	*8.9702*	*−46.3803*	*0.1036*	*−4.7887*	
LNF2	20.4382	18.370	−23.824	0.1448	−3.4493	5
	9.0736	*9.0736*	*−47.3634*	*0.1021*	*−4.8624*	
LNF3	15.984	14.249	−29.4120	0.1304	−3.8338	5
	9.0080	*9.0080*	*−46.4638*	*0.1035*	*−4.7919*	
LNF4	15.849	14.219	−29.7070	0.1297	−3.8530	5
	9.0890	*9.0890*	*−46.2793*	*0.1038*	*−4.7753*	
HT						
LNS	21.091	19.769	−22.5432	0.1489	−3.3566	5
	9.4415	*9.4415*	*−46.9805*	*0.1035*	*−4.8022*	
LNF1	21.672	20.046	−22.2365	0.1498	−3.3331	5
	9.5269	*9.5269*	*−46.7829*	*0.1039*	*−4.7885*	
LNF2	23.074	21.052	−21.1735	0.1536	−3.2521	5
	9.5392	*9.5392*	*−46.7340*	*0.1040*	*−4.7855*	
LNF3	29.796	25.271	−17.6496	0.1683	−2.9706	5
	9.5666	*9.5666*	*−46.6935*	*0.1042*	*−4.7798*	
LNF4	24.043	21.728	−20.7876	0.1549	−3.2206	5
	9.2045	9.2045	*−48.1036*	0.1013	*−4.8975*	
NG						
LNS	15.788	14.018	−32.3786	0.1235	−4.0010	5
	9.5375	*9.5375*	*−46.8690*	*0.1038*	*−4.7949*	
LNF1	18.934	16.187	−27.8293	0.1337	−3.7215	5
	9.3295	*9.3295*	*−47.5860*	*0.1023*	*−4.8576*	
LNF2	23.364	20.923	−20.9780	0.1535	−3.2205	5
	9.1757	*9.1757*	*−47.1732*	*0.1028*	*−4.8296*	
LNF3	20.586	18.2140	−24.8640	0.1417	−3.5244	5
	9.2718	*9.2718*	*−47.7026*	*0.1019*	*−4.8803*	
LNF4	21.818	19.8910	−20.9785	0.1518	−3.1858	5
	8.8876	*8.8876*	*−46.0276*	*0.1038*	*−4.7703*	

*Denotes 5% significance level. Figures shown in standard format are test statistics. Italicised and underlined figures indicate the critical values at the 5% significance level. m shows the number of breaks determined by the CS test. WTI: West Texas Intermediate, ZA: Zivot and Andrews, ADF: Augmented Dickey–Fuller, NG: Natural gas

obtained in this case are different from what the GH and HJ tests indicate. In other words, it shows that spot and futures prices are not cointegrated for seven cases out of 16, constituting nearly 44% of all the pairs of spot and futures prices. One possible reason why Maki (2012) test presents such different results may be that this test should be estimated using a trimming region of 5%, as proposed by Maki (2012). However, up to now, as it is a more common approach in the literature, the trimming region is set to be 0.15 in all cases. Therefore, following Maki (2012), we also estimate the test with a trimming region of 0.05 to examine the robustness of the test; Table 4 presents the results. Here, we obtain similar findings to the GH and HJ tests. In other words, the Maki (2012) test shows that all pairs of spot and futures prices of energy commodities are cointegrated at the 5% or higher significance level with only two exceptions for Brent, where spot and futures prices at 3 and 4 months maturities are found not to be cointegrated. These results imply that Maki's (2012) test findings may be sensitive to trimming value and spot and futures prices may not be cointegrated, at least in some cases, when it is allowed for five regime shifts.

Because of these mixed results from the Maki (2012) test and that the relevant studies have thus far focused less on the impact of regime shifts on market efficiency, we decide to concentrate on the results from the GH and HJ tests, and leave the potential impact of more than two regime shifts (i.e. five regime shifts) on market

Table 3: Gregory and Hansen (1996) and Hatemi-J (2008) cointegration tests results

Model	Gregory and Hansen (1996)			Hatemi-J (2008)		
	ADF*	Z^*_t	Z^*_α	ADF*	Z^*_t	Z^*_α
WTI						
LS-LF1	−7.731*	−28.25*	−788.70*	−8.05*	−28.77*	−802.6*
	(0.406)	(0.416)	(0.416)	(0.296, 0.320)	(0.358, 0.517)	(0.358, 0.517)
LS-LF2	−7.23*	−10.30*	−177.9*	−7.82*	−11.3*	−208.2*
	(0.402)	(0.402)	(0.402)	(0.154, 0.257)	(0.362, 0.515)	(0.362, 0.515)
LS-LF3	−7.23*	−8.03*	−115.8*	−8.75*	−8.77*	−135.8*
	(0.673)	(0.402)	(0.402)	(0.397, 0.522)	(0.389, 0.512)	(0.389, 0.512)
LS-LF4	−7.546*	−7.366*	−99.68*	−8.23*	−7.96*	−115.5*
	(0.402)	(0.410)	(0.410)	(0.401, 0.512)	(0.389, 0.511)	(0.389, 0.511)
Brent						
LS-LF1	−12.38*	−16.92*	−408.0*	−12.98*	−17.71*	−436.9*
	(0.388)	(0.388)	(0.388)	(0.388, 0.502)	(0.388, 0.507)	(0.388, 0.507)
LS-LF2	−9.82*	−12.20*	−241.3*	−10.61*	−13.28*	−279.4*
	(0.388)	(0.388)	(0.388)	(0.388, 0.501)	(0.388, 0.503)	(0.388, 0.503)
LS-LF3	−8.50*	−10.05*	−173.8*	−9.39*	−11.3*	−213.4*
	(0.388)	(0.388)	(0.388)	(0.388, 0.508)	(0.388, 0.507)	(0.388, 0.504)
LS-LF4	−8.04*	−8.83*	−139.3*	−8.620*	−10.02*	−174.3*
	(0.671)	(0.412)	(0.412)	(0.388, 0.508)	(0.388, 0.507)	(0.388, 0.507)
HT						
LS-LF1	−8.54*	−9.07*	−147.5*	−8.86*	−9.36*	−156.12*
	(0.302)	(0.300)	(0.300)	(0.150, 0.151)	(0.151, 0.153)	(0.151, 0.153)
LS-LF2	−6.58*	−7.52*	−105.2*	−7.75*	−7.93*	−115.99*
	(0.300)	(0.300)	(0.300)	(0.153, 0.180)	(0.153, 0.155)	(0.153, 0.155)
LS-LF3	−6.47*	−6.61*	−83.30*	−7.04*	−7.00*	−92.76*
	(0.332)	(0.300)	(0.300)	(0.150, 0.275)	(0.153, 0.277)	(0.153, 0.277)
LS-LF4	−5.63*	−6.48*	−79.79*	−6.456**	−6.87*	−90.012**
	(0.426)	(0.427)	(0.427)	(0.150, 0.646)	(0.153, 0.277)	(0.153, 0.277)
NG						
LS-LF1	−11.79*	−11.79*	−236.4*	−12.2*	−12.2*	−252.1*
	(0.279)	(0.279)	(0.279)	(0.317, 0.585)	(0.317, 0.585)	(0.316, 0.585)
LS-LF2	−9.04*	−9.14*	−151.4*	−9.39*	−9.54*	−164.2*
	(0.279)	(0.279)	(0.279)	(0.379, 0.585)	(0.378, 0.584)	(0.378, 0.584)
LS-LF3	−7.38*	−7.47*	−104.4*	−7.62*	−7.78*	−113.06*
	(0.279)	(0.279)	(0.279)	(0.371, 0.585)	(0.371, 0.584)	(0.371, 0.584)
LS-LF4	−6.88*	−6.76*	−86.79*	−7.11*	−6.97*	−92.03*
	(0.278)	(0.279)	(0.279)	(0.266, 0.267)	(0.371, 0.577)	(0.371, 0.577)

The GH test critical values are from Table 1 of GH (1996. p. 109), and HJ test critical values are from Table 1 of HJ (2008. p. 501). Numbers in parentheses denote the break points. 15% trimming region is used for the GH and HJ tests. * and ** denote the rejection of null hypothesis of no cointegration at the 1% and 5% significance levels, respectively. LF1, LF2, LF3, and LF4 are the futures prices at 1, 2, 3, and 4 months maturity, respectively. WTI: West Texas Intermediate, ZA: Zivot and Andrews, ADF: Augmented Dickey–Fuller, NG: Natural gas

efficiency for future studies. However, it is also worth noting that the main contribution of using the Maki (2012) test in our study is that contrary to the general findings in the relevant literature that indicate that spot and futures prices are cointegrated (e.g., Switzer and El-Khoury, 2007; Tse, 1995), in fact, the results from the Maki (2012) test show that they may not be cointegrated at least in some cases. This finding is also consistent with the recent findings by Wang and Wu (2013) who investigate the cointegration relationship between the spot and futures prices using the nonlinear threshold vector error correction model. Therefore, future studies can investigate whether other spot and futures prices are cointegrated by employing the Maki (2012) test. We believe that such an approach may provide further evidence to the relevant literature.

Turning to the GH and HJ tests results, as indicated before, both tests show that all pairs of spot and futures prices are cointegrated in all cases. However, it is worth noting that the cointegration relationship is just a necessary condition for unbiasedness hypothesis. Additionally, for futures markets to be unbiased predictors of future spot prices, $\alpha_0 = 0$ and $\beta_0 = 1$ restrictions should also hold. However, as stated previously, since ensuring $\beta_0 = 1$ restriction is more important in terms of market efficiency, we also test $\alpha_0 = 0$ and $\beta_0 = 1$ restrictions separately. Thus, we first estimate Equation (1) that does not consider the structural breaks and check whether relevant restrictions hold. Then, to allow for the potential impact of one and two regime shifts on parameter estimates and market efficiency, based on the GH and HJ cointegration test results, we estimate Equations (2) and (3) considering one and two breaks, and we further analyse whether the relevant restrictions hold[5]. Following Abosedra

5 More specifically, after estimating equation (1), Ho: $\alpha_0 = 0$ and Ho: $\beta_0 = 1$ hypotheses are tested both jointly and separately; similarly, after estimating equation (2), Ho: $\alpha_0 + \alpha_1 = 0$ and Ho: $\beta_0 + \beta_1 = 1$ hypotheses are tested both jointly and separately, and lastly after estimating equation (3), Ho: $\alpha_0 + \alpha_1 + \alpha_2 = 0$ and Ho: $\beta_0 + \beta_1 + \beta_2 = 1$ hypotheses are tested both jointly and separately. Additionally, for simplicity, throughout the paper, α parameter is used to represent both the $\alpha_0 + \alpha_1$ in equation (2) and $\alpha_0 + \alpha_1 + \alpha_2$ in equation (3). Accordingly, β parameter is used to represent both the $\beta_0 + \beta_1$ in equation (2) and $\beta_0 + \beta_1 + \beta_2$ in equation (3).

Table 4: The Maki (2012) cointegration test results

Model	Maki test statistic		Break points when triminning value is 0.05				
	Trimming value						
	0.15	0.05	T_{B1}	T_{B2}	T_{B3}	T_{B4}	T_{B5}
WTI							
LS-LF1	−7.94*	−7.94*	0.357	0.413	0.669	0.720	0.832
LS-LF2	−7.36*	−8.38*	0.140	0.240	0.417	0.669	0.722
LS-LF3	−6.29	−8.27*	0.140	0.250	0.417	0.556	0.669
LS-LF4	−5.84	−7.14*	0.082	0.347	0.417	0.556	0.669
Brent							
LS-LF1	−5.52	−7.10*	0.114	0.248	0.390	0.659	0.738
LS-LF2	−5.27	−7.07*	0.114	0.248	0.390	0.659	0.824
LS-LF3	−4.94	−5.468	0.114	0.390	0.460	0.679	0.845
LS-LF4	−4.94	−5.461	0.082	0.248	0.390	0.617	0.845
HT							
LS-LF1	−6.40*	−7.10*	0.074	0.154	0.279	0.403	0.947
LS-LF2	−6.19	−7.03*	0.074	0.130	0.196	0.279	0.442
LS-LF3	−6.41*	−7.01*	0.074	0.200	0.291	0.349	0.420
LS-LF4	−6.72*	−7.39*	0.074	0.275	0.348	0.442	0.798
NG							
LS-LF1	−8.17*	−8.17*	0.176	0.227	0.280	0.397	0.731
LS-LF2	−7.62*	−7.82*	0.280	0.397	0.459	0.523	0.730
LS-LF3	−7.21*	−7.52*	0.135	0.224	0.280	0.397	0.458
LS-LF4	−6.91*	−7.13*	0.130	0.223	0.280	0.468	0.724

The Maki (2012) test critical values are from Table 1 of Maki (2012. p. 2013). *Denote the rejection of null hypothesis of no cointegration at the 5% significance level. LF1, LF2, LF3, and LF4 are the futures prices at 1, 2, 3, and 4 months maturity, respectively. WTI: West Texas Intermediate, ZA: Zivot and Andrews, ADF: Augmented Dickey–Fuller, NG: Natural gas

and Baghestani (2004), Hatemi-J (2008), Narayan and Narayan (2010), and Kanjilal and Ghosh (2013), we estimate Equations (1), (2), and (3) with ordinary least squares; Table 5 presents the results[6]. First, if we examine the parameter estimation results, both adjusted R^2 and akaike information criterion (AIC) values in all cases show that the models with one and/or two structural breaks are more appropriate than the model without structural breaks. The results also indicate that among the alternative structural break models, nearly in all cases, the model with two regime shifts should be preferred to the model with one regime shift.

Second, we analyse the market efficiency; Tables 6 and 7 illustrates the results. Starting with Equation (1), the results show that the null hypothesis of $\alpha_0 = 0$ and $\beta_0 = 1$ is rejected jointly at the 5% significance level in all cases, implying that none of the energy futures markets are unbiased predictors of future spot prices. Besides, when the relevant restrictions are tested separately, the results show that the null hypothesis of $\alpha_0 = 0$ is rejected in all cases for WTI, Brent, and NG, whereas it holds for HT in all cases. This implies that there is non-zero risk premium in all cases except for HT. Further, we see that $\beta_0 = 1$ restriction is rejected for WTI and HT futures markets in all cases, whereas it holds for Brent futures market in the case of nearest contract and for NG futures market in all cases at the 5% significance level. These results reveal that although none of the energy futures markets is unbiased predictors of future spot prices, Brent and NG futures markets are found to be efficient because they hold $\beta_0 = 1$ restriction. However, it is worth noting that while efficiency is ensured for NG in all cases, it is valid for Brent only for the nearest futures contract.

As for the impact of structural breaks, starting with Equation (2) which considers only one regime shift, the results show that the null hypothesis of $\alpha_0 = 0$ and $\beta_0 = 1$ is again rejected jointly at the 5% significance level in all cases. Besides, when the relevant restrictions are tested separately, the results show that the null hypothesis of $\alpha_0 = 0$ is rejected in all cases including HT. Moreover, in this case, $\beta_0=1$ restriction hold for HT futures market in the case of nearest three contracts and for NG in all cases, whereas it is rejected for Brent and WTI futures market for all four futures contracts at the 5% significance level. Therefore, the existence of structural break has an important impact on testing market efficiency. From the results of Equation (3), which allows two regime shifts, we see that the null hypothesis of $\alpha_0=0$ and $\beta_0 = 1$ is again rejected jointly at the 5% significance level in all cases, meaning that none of the energy futures markets are unbiased predictors of future spot prices. Besides, when the relevant restrictions are tested separately, the findings indicate that the null hypothesis of $\alpha_0 = 0$ is rejected in all cases, except for HT futures contracts at 3 and 4 months maturity. Further, we see that $\beta_0 = 1$ restriction does not hold for Brent and WTI futures market in any cases, whereas the restriction holds for HT futures market in the case of nearest contract and for NG futures market in all cases at the 5% significance level. Therefore, Equations (2) and (3) generally provide similar results in terms of market efficiency because both equations show that HT and NG futures markets are efficient, whereas Brent and WTI futures markets are not. However, it is also worth noting that the results also imply that the number of regime shifts may have an impact on testing market efficiency, although the impact is not as significant as that of the existence of structural breaks.

Lastly, it is also worth noting that nearly in all cases, AIC (together with adjusted R^2) indicates that, among the alternative three models, the most appropriate model is one with two-regime shifts, which is represented by Equation (3). Additionally, as discussed in the introduction section, it is more likely for energy commodities

6 Following HJ, while estimating Equations (2) and (3), we consider the break points determined by Z_t test statistic.

Table 5: Parameter estimation results

Model	α_0	α_1	α_2	β_0	β_1	β_2	R^2adjusted	AIC
Without structural break								
WTI								
LS-LF1	0.0056*	-	-	0.9984*	-	-	0.9998	−6.8577
LS-LF2	0.0398*	-	-	0.9892*	-	-	0.9982	−4.6323
LS-LF3	0.0824*	-	-	0.9784*	-	-	0.9958	−3.7742
LS-LF4	0.1275*	-	-	0.9674*	-	-	0.9933	−3.2925
Brent								
LS-LF1	−0.0128*	-	-	1.002*	-	-	0.9984	−4.5435
LS-LF2	0.0126	-	-	0.9955*	-	-	0.9969	−3.8661
LS-LF3	0.0400*	-	-	0.9888*	-	-	0.9949	−3.3712
LS-LF4	0.0725*	-	-	0.9812*	-	-	0.9928	−3.0103
HT								
LS-LF1	−0.0012	-	-	0.9925*	-	-	0.9978	−4.2802
LS-LF2	−0.0005	-	-	0.9856*	-	-	0.9945	−3.3573
LS-LF3	0.0015	-	-	0.9782*	-	-	0.9910	−2.8565
LS-LF4	0.0050**	-	-	0.9708*	-	-	0.9870	−2.4888
NG								
LS-LF1	−0.0183*	-	-	0.9995*	-	-	0.9852	−2.9977
LS-LF2	−0.0356*	-	-	0.9938*	-	-	0.9619	−2.0543
LS-LF3	−0.0417*	-	-	0.9853*	-	-	0.9343	−1.5073
LS-LF4	−0.0468*	-	-	0.9799*	-	-	0.9059	−1.1492
With one structural break								
WTI								
LS-LF1	−0.0036	−0.0275*	-	1.002*	0.0052**	-	0.9998	−6.9135
LS-LF2	−0.0587*	−0.1676*	-	1.021*	0.0278*	-	0.9988	−5.0797
LS-LF3	−0.0983*	−0.2554*	-	1.036*	0.0396*	-	0.9975	−4.2898
LS-LF4	−0.1032*	−0.3824*	-	1.042*	0.0631*	-	0.9962	−3.8519
Brent								
LS-LF1	−0.1322*	−0.0040	-	1.0403*	−0.011**	-	0.9987	−4.7312
LS-LF2	−0.1736*	−0.0799*	-	1.0564*	−0.0020	-	0.9979	−4.2254
LS-LF3	−0.2207*	−0.1295*	-	1.0744*	0.0010	-	0.9968	−3.8255
LS-LF4	−0.1425*	−0.3832*	-	1.0525*	0.0614*	-	0.9955	−3.4925
HT								
LS-LF1	0.0256*	−0.0351*	-	1.0454*	−0.0450*	-	0.9981	−4.3860
LS-LF2	0.0488*	−0.0662*	-	1.0820*	−0.0791*	-	0.9953	−3.4928
LS-LF3	0.0721*	−0.0906*	-	1.1184*	−0.1203*	-	0.9923	−3.0034
LS-LF4	0.0380*	−0.1247*	-	1.0416*	0.0250	-	0.9888	−2.6366
NG								
LS-LF1	−0.0371*	0.0140	-	1.0215*	−0.0209**	-	0.9854	−3.0071
LS-LF2	−0.0926*	0.0542*	-	1.0529*	−0.0607*	-	0.9629	−2.0780
LS-LF3	−0.1690*	0.1224*	-	1.1110*	−0.1288*	-	0.9380	−1.5638
LS-LF4	−0.2366*	0.1729*	-	1.1652*	−0.1837*	-	0.9132	−1.2271
With two structural breaks								
WTI								
LS-LF1	−0.0127*	0.0228	−0.0461*	1.0044*	−0.0073**	0.0106*	0.9998	−6.9167
LS-LF2	−0.1206*	0.2291*	−0.4004*	1.0408*	−0.0705*	0.0931*	0.9990	−5.1914
LS-LF3	−0.1470*	0.2604*	−0.5817*	1.0520*	−0.0850*	0.1344*	0.9978	−4.3941
LS-LF4	−0.1832*	0.3932*	−0.8209*	1.0671*	−0.1247*	0.1899*	0.9966	−3.9652
Brent								
LS-LF1	−0.1322*	0.0728	−0.1261*	1.0403*	−0.0286**	0.0279**	0.9987	−4.7508
LS-LF2	−0.1736*	0.1352**	−0.3142*	1.0565*	−0.0521*	0.0716*	0.9979	−4.2767
LS-LF3	−0.2207*	0.2297*	−0.5179*	1.0744*	−0.0832*	0.1188*	0.9971	−3.9106
LS-LF4	−0.2589*	0.3504*	−0.7351*	1.0900*	−0.1199*	0.1692*	0.9960	−3.6216
HT								
LS-LF1	0.0332*	−0.0320	−0.0067	1.0443*	−0.2040	0.1560	0.9980	−4.3771
LS-LF2	0.0635*	−0.0218	−0.0503	1.0788*	0.0766	−0.1621	0.9952	−3.4701
LS-LF3	0.0934*	−0.0449*	−0.0501*	1.1186*	0.0396	−0.1794*	0.9923	−3.0091
LS-LF4	0.1266*	−0.0693*	−0.0544*	1.1648*	0.0048	−0.2001*	0.9889	−2.6460
NG								
LS-LF1	−0.0439*	0.0128	−0.0004	1.0274*	−0.0272	0.0103	0.9856	−3.0257
LS-LF2	−0.0812*	0.0318	−0.0156	1.0398*	−0.0512	0.0226	0.9636	−2.0949
LS-LF3	−0.1408*	0.1265	−0.0812	1.0821*	−0.1276*	0.0594	0.9401	−1.5957
LS-LF4	−0.1835*	0.1652	−0.0976	1.1143*	−0.1661*	0.0628	0.9162	−1.2593

*Denotes 5% significance level. WTI: West Texas Intermediate, ZA: Zivot and Andrews, ADF: Augmented Dickey–Fuller, NG: Natural gas, AIC: Akaike information criterion

Table 6: Tests of market efficiency hypothesis

Model	OLS					
	Without structural break			With one endogenous structural break		
	$\alpha_0=0, \beta_0=1$	$\alpha_0=0$	$\beta_0=1$	$\alpha_0=0, \beta_0=1$	$\alpha_0=0$	$\beta_0=1$
WTI						
LS-LF1	6.44* (0.0020)	8.29* (0.0040)	9.96* (0.0020)	24.76* (0.000)	25.67* (0.000)	22.78* (0.000)
LS-LF2	32.51* (0.000)	45.24* (0.000)	53.27* (0.000)	200.4* (0.000)	217.7* (0.000)	194.6* (0.000)
LS-LF3	48.25* (0.000)	83.98* (0.000)	91.82* (0.000)	221.5* (0.000)	227.3* (0.000)	202.6* (0.000)
LS-LF4	65.98* (0.000)	126.6* (0.000)	131.6* (0.000)	236.0* (0.000)	246.5* (0.000)	221.4* (0.000)
Brent						
LS-LF1	26.75* (0.000)	5.17** (0.023)	1.322 (0.2510)	80.05* (0.000)	86.29* (0.000)	74.96* (0.000)
LS-LF2	10.91* (0.000)	2.543 (0.1110)	5.15** (0.023)	132.4* (0.000)	170.8* (0.000)	153.0* (0.000)
LS-LF3	12.69* (0.000)	15.79* (0.000)	19.51* (0.000)	158.6* (0.000)	208.9* (0.000)	188.3* (0.000)
LS-LF4	19.87* (0.000)	36.61* (0.000)	39.16* (0.000)	185.2* (0.000)	264.5* (0.000)	243.5* (0.000)
HT						
LS-LF1	17.74* (0.000)	0.977 (0.3230)	20.51* (0.000)	30.91* (0.000)	16.54* (0.000)	0.036 (0.8480)
LS-LF2	21.69* (0.000)	0.089 (0.7640)	30.13* (0.000)	37.00* (0.000)	22.35* (0.000)	0.338 (0.5610)
LS-LF3	26.29* (0.000)	0.375 (0.5400)	42.46* (0.000)	37.17* (0.000)	15.50* (0.000)	0.112 (0.7380)
LS-LF4	29.36* (0.000)	2.974 (0.0850)	53.31* (0.000)	63.37* (0.000)	66.14* (0.000)	30.9* (0.0000)
NG						
LS-LF1	48.62* (0.000)	6.79* (0.0090)	0.014 (0.9070)	47.58* (0.000)	5.438* (0.000)	0.009 (0.9230)
LS-LF2	106.6* (0.000)	9.60* (0.0020)	0.754 (0.3850)	103.4* (0.000)	5.695* (0.017)	0.718 (0.3970)
LS-LF3	128.2* (0.000)	7.367* (0.007)	2.466 (0.1160)	138.2* (0.000)	4.804* (0.028)	2.174 (0.1410)
LS-LF4	132.2* (0.000)	6.27** (0.013)	3.136 (0.0770)	151.9* (0.000)	6.096* (0.014)	1.633 (0.2020)

OLS denotes ordinary least squares. Relevant null hypotheses are examined with the Wald test. Numbers in parentheses are P values. * and ** denote1% and 5% significance levels, respectively. LF1, LF2, LF3, and LF4 are the futures prices at 1, 2, 3, and 4 months maturity, respectively. WTI: West Texas Intermediate, ZA: Zivot and Andrews, ADF: Augmented Dickey–Fuller, NG: Natural gas

Table 7: Tests of market efficiency hypothesis

Model	With two endogenous structural breaks		
	$\alpha_0=0, \beta_0=1$	$\alpha_0=0$	$\beta_0=1$
WTI			
LS-LF1	22.59* (0.000)	27.38* (0.000)	24.84* (0.000)
LS-LF2	215.1* (0.000)	287.8* (0.000)	265.7* (0.000)
LS-LF3	245.2* (0.000)	317.2* (0.000)	292.7* (0.000)
LS-LF4	253.7* (0.000)	322.1* (0.000)	297.4* (0.000)
Brent			
LS-LF1	76.6* (0.000)	102.5* (0.000)	93.32* (0.000)
LS-LF2	138.4* (0.000)	214.8* (0.000)	200.5* (0.000)
LS-LF3	181.2* (0.000)	289.5* (0.000)	271.8* (0.000)
LS-LF4	202.5* (0.000)	327.5* (0.000)	308.7* (0.000)
HT			
LS-LF1	25.88* (0.000)	13.89* (0.000)	3.428 (0.0640)
LS-LF2	28.68* (0.000)	12.74* (0.000)	4.257* (0.039)
LS-LF3	32.74* (0.000)	14.27 (0.7056)	16.04* (0.000)
LS-LF4	31.72* (0.000)	0.3300 (0.564)	23.21* (0.000)
NG			
LS-LF1	21.76* (0.000)	8.812* (0.003)	1.909 (0.167)
LS-LF2	52.49* (0.000)	11.47* (0.001)	0.842 (0.359)
LS-LF3	75.72* (0.000)	14.12* (0.000)	0.752 (0.386)
LS-LF4	97.46* (0.000)	14.34* (0.000)	0.335 (0.563)

OLS denotes ordinary least squares. Relevant null hypotheses are examined with the Wald test. Numbers in parentheses are P values. * and ** denote 1% and 5% significance levels, respectively. LF1, LF2, LF3, and LF4 are the futures prices at 1, 2, 3, and 4 months maturity, respectively. WTI: West Texas Intermediate, ZA: Zivot and Andrews, ADF: Augmented Dickey–Fuller, NG: Natural gas

to have multiple structural breaks. We hence give more weight to the results of Equation (3). Therefore, we come to the following conclusion: (i) Because $\alpha = 0$ and $\beta_0 = 1$ hypothesis is rejected jointly in all cases, none of the energy futures markets examined are unbiased estimator of future spot prices. (ii) However, because $\beta_0 = 1$ restriction holds for NG in the case of all futures contracts and for HT in the case of nearest futures contract, NG and HT futures markets are efficient markets.

4. CONCLUDING REMARKS AND POLICY IMPLICATIONS

This study examines the WTI, Brent, heating oil, and NG futures markets' efficiency, using spot and futures prices at 1, 2, 3, and 4 months maturity based on the multiple structural breaks. First, three different unit root tests allowing no break, one endogenous break, and five endogenous breaks are used to examine the stationary properties of the relevant energy commodities. Then, the Gregory and Hansen (1996), Hatemi-J (2008), and Maki (2012) cointegration tests allowing one, two, and five endogenous breaks are used, respectively. Besides, the cointegrating coefficients are estimated with a similar method used by Hatemi-J (2008) that considers the impact of structural breaks on parameter estimates.

Our main findings are as follows. First, we find that spot and futures prices at four different maturities are not stationary at their level form even if five endogenous breaks are allowed in the level and slope of time trend. This means that a shock to these energy commodities may have a permanent effect. Second, we find that all pairs of spot and futures prices have a cointegration relationship even if up to two regime shifts are allowed. This implies that spot and futures prices have a common stochastic trend. That is, they are driven by the same main factors, and the length of futures contracts does not have a noticeably different impaction the cointegration relationship. However, based on the Maki (2012) cointegration test, when the five regime shifts are considered, the mixed results are obtained because of the sensitivity of the Maki (2012) test to the trimming value. Therefore, we think that future studies could investigate whether other spot and futures prices are cointegrated by employing the Maki (2012) test. Such an approach can provide further evidence to the relevant literature. Third, it is important to consider the structural breaks when determining whether energy futures markets are efficient. In this regard, when we consider the impact of the two structural breaks, we find that WTI and Brent futures markets are not efficient markets based on Fama's (1970) hypothesis, whereas NG and heating oil markets are found to be efficient. However, although efficiency is ensured for NG in the case of all four futures contracts, it is valid for heating oil market only in the case of nearest futures contract.

These results imply that WTI and Brent futures markets can provide consistent abnormal profit for traders, and are not unbiased predictors of future spot prices; thus, these should not be used by economic agents to forecast spot prices. Therefore, regulators should try to improve the information flows and reduce possible market manipulation in these markets (Maslyuk and Smyth, 2009; Stout, 1995). Contrarily, although NG and heating oil futures markets are not unbiased predictors of spot prices either, we find that they are efficient because they hold $\beta_0 = 1$ restriction. However, it is worth noting that while efficiency is ensured for NG in the case of all four futures contracts, it is valid for heating oil market only in the case of nearest futures contract. This means that risk managers can use all four futures contracts to hedge the market risk of NG, but only the nearest futures contract to hedge the market risk of heating oil.

5. ACKNOWLEDGMENTS

This research did not receive any specific grant from any funding agencies in the public, commercial, or not-for-profit sectors.

REFERENCES

Abosedra, S., Baghestani, H. (2004), On the predictive accuracy of crude oil futures prices. Energy Policy, 32, 1389-1393.

Arouri, M.E.H., Hammoudeh, S., Lahiani, A., Nguyen, D.K. (2013), On the short-and long-run efficiency of energy and precious. Energy Economics, 40, 832-844.

Askari, H., Krichene, N. (2008), Oil price dynamics (2002-2006). Energy Economics, 30, 2134-2153.

Bai, J., Perron, P. (1998), Estimating and testing linear models with multiple structural changes. Econometrica, 66, 47-78.

Bai, J., Perron, P. (2003), Computation and analysis of multiple structural change models. Journal of Applied Econometrics, 18, 1-22.

Beck, S.E. (1994), Cointegration and market efficiency in commodities futures markets. Applied Economics, 26(3), 249-257.

Çağlı, E.F, Mandaci, P.E. (2013), The long-run relationship between the spot and futures markets under multiple regime-shifts: Evidence from Turkish derivatives exchange. Expert Systems with Applications, 40, 4206-4212.

Carrion-I-Silvestre, J.L., Kim, D., Perron, P. (2009), GLS-based unit root tests with multiple structural breaks under both the null and alternative hypotheses. Econometric Theory, 25(6), 1754-1792.

Chen, P.F., Lee, C.C., Zeng, J.H. (2014), The relationship between spot and futures oil price: Do structural break matter? Energy Economics, 43, 206-217.

Chin, M.D., Le Blanc, M., Coibion, O. (2005), The Predictive Content of Energy Futures: An Update on Petroleum, Natural Gas, Heating Oil and Gasoline, National Bureau of Economic Research NBER, Working Paper, No. 11033.

Crowder, W.J., Hamid, A. (1993), A co-integration test for oil futures market efficiency. Journal of Futures Markets, 13, 933-941.

Dickey, D., Fuller, W. (1979), Distribution of the estimators for autoregressive time series with unit root. Journal of American Statistical Association, 74, 427-431.

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

Fan, Y., Xu, J.H. (2011), What has driven oil prices since 2000? A structural change perspective. Energy Economics, 33, 1082-1094.

Garbade, K.D., Silber, W.L. (1983), Price movements and price discovery in futures and cash markets. The Review of Economics and Statistics, 65(2), 289-297.

Gebre-Mariam, Y.K. (2011), Testing for unit roots, causality, cointegration, and efficiency: The case of the northwest US natural gas market. Energy, 36, 3489-3500.

Gregory, A.W., Hansen, B.E. (1996), Residual-based tests for cointegration in models with regime shifts. Journal of Econometrics, 70, 99-126.

Hatemi, J.A. (2008), Tests for cointegration with two unknown regime shifts with an application to financial market integration. Emprical Economics, 35, 497-505.

Johansen, S. (1988), Statistical analysis of cointegration vectors. Journal of Economics Dynamics and Control, 12, 231-254.

Kanjilal, K., Ghosh, S. (2013), Environmental Kuznet's curve for India: Evidence from tests for cointegration with unknown structural breaks. Energy Policy, 56, 509-515.

Kapetanious, G. (2005), Unit root testing against the alternative hypothesis of up to m structural breaks. Journal of Time Series Analysis, 26, 123-133.

Kawamoto, K., Hamori, S. (2011), Market efficiency among futures with different maturities: Evidence from the crude oil futures market. Journal of Futures Markets, 31, 487-501.

Lean, H.H., McAleer, M., Wong, W.K. (2010), Market efficiency of oil spot and futures: A mean-variance and stochastic dominance approach. Energy Economics, 32, 979-986.

Lee, C.C., Lee, J.D. (2009), Energy prices, multiple structural breaks, and efficient market hypothesis. Applied Energy, 86, 466-479.

Lee, C.C., Zeng, J.H. (2011), Revisiting the relationship between spot and future oil prices: Evidence from quantile cointegration regression. Energy Economics, 33, 924-935.

Maki, D. (2012), Tests for cointegration allowing for an unknown number of breaks. Economic Modelling, 29, 2011-2015.

Maslyuk, S., Smyth, R. (2009), Cointegration between oil spot and future prices of the same and different grades in the presence of structural change. Energy Policy, 37, 1687-1693.

McKenzi, A.M., Holt, M.T. (2002), Market efficiency in agricultural futures. Applied Economics, 34, 1519-1532.

Moosa, I.A., Al-Loughani, N.E. (1994), Unbiasedness and time varying risk premia in the crude oil futures market. Energy Economics, 16(2), 99-105.

Narayan, P.K, Narayan, S. (2010), Modeling the impact of oil prices on Vietnam's stock prices. Applied Energy, 87, 356-361.

Narayan, P.K., Liu, R. (2011), Are shocks to commodity price persistence? Applied Energy, 88, 409-416.

Narayan, P.K., Narayan, S., Zheng, X. (2010), Gold and oil future markets: Are markets efficient. Applied Energy, 87, 3299-3303.

Ng, S., Perron, P. (2003), Lag length selection and construction of unit root tests with good size and power. Econometrica, 69(6), 1519.

Noguera, J. (2013), Oil prices: Breaks and trends. Energy Economics, 37, 60-67.

Ozdemir, Z.A., Gokmenoglu, K., Ekinci, C. (2013), Persistence in crude oil spot and futures prices. Energy, 59, 29-37.

Peroni, E., McNown, R. (1998), Noninformative and informative tests of efficiency in three energy futures market. Journal of Futures Markets, 18, 939-964.

Perron, P. (1989), The great crash, the oil price shocks, and the unit root hypothesis. Econometrica, 57, 1361-1401.

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

Shambora, W.E., Rossiter, R. (2007), Are there exploitable inefficiencies in the futures market for oil? Energy Economics, 29, 18-27.

Stout, L.A. (1995), Are stock markets costly casinos? Disagreement, market failure and securities regulation. Virginia Law Review, 81, 611-712.

Switzer, L.N., El-Khoury, M. (2007), Extreme volatility, speculative efficiency and hedging effectiveness on the oil future markets. Journal of Futures Markets, 27, 61-84.

Tang, B.J., Shen, C., Gao, C. (2013), The efficiency analysis of European CO_2 futures market. Applied Energy, 112, 1544-1547.

Tse, Y.K. (1995), Lead-lag relationship between spot index and futures price of the Nikkei stock average. Journal of Forecasting, 14(7), 553-563.

Wang, H.H., Ke, B. (2005), Efficiency tests of agricultural commodity futures markets in China. Australian Journal of Agricultural and Resource Economics, 49, 125-141.

Wang, Y., Wu, C. (2013), Are crude oil spot and futures prices cointegrated? Not always! Economic Modelling, 33, 641-650.

Zhang, Y.J., Wang, Z.Y. (2013), Investigating the price discovery and risk transfer functions in the crude oil and gasoline futures markets: Some empirical evidence. Applied Energy, 104, 220-228.

Zivot, E., Andrews, D.W.K. (1992), Further evidence on the great crash, the oil price shocks, and unit root hypothesis. Journal of Business and Economic Statistics, 10, 251-270.