Threshold Cointegration and Dynamics of Crude Oil Futures Volatility and Financial Speculation*

Nicholas Ruei-Lin Lee**, Hsiang-Hui Chu***, Jung-Fang Liu****, and Yih-Bey Lin*****

Abstract

This paper investigates threshold cointegration and dynamics of New York Mercantile Exchange (NYMEX) crude oil futures volatility and speculation in the framework of a threshold vector error correction model (TVECM). Two regimes are determined by the model and divided into the usual and unusual regime. We use the ratio of volume over open interest as speculation variable while we consider two volatilities measures, absolute return and high-low

* The authors are grateful to two anonymous reviewers for their valuable suggestions, leading to substantial improvements for this paper. Any shortcomings and errors are the author’s responsibility.

** Assistant Professor, Department of Finance, Chaoyang University of Technology, Taiwan.

*** Assistant Professor, Department of Banking and Finance, National Chi Nan University, Taiwan.

**** Lecturer, Department of Public Finance, Ling Tung University, Taichung, Taiwan.

***** Professor, Department of Finance, Chaoyang University of Technology, Taiwan.

Corresponding Author. Email: ybblin@cyut.edu.tw, TEL: +886-4-2332-3000 ext7092.

DOI: 10.3966/054696002014120096001
volatilities. Our findings present different and strong asymmetric error correction effects of volatility and speculation in the speed of adjustment to the long-run equilibrium in two regimes. In addition of speculation leading volatility based on former measure, speculation following volatility is found in the usual regime, suggesting relative market efficient. Bidirectional causalities of speculation and different volatilities measures in the unusual regime suggest that market instability mitigated in the unusual regime. In sum, crude oil futures markets seem to work well with the existence of threshold cointegration.

**Keywords:** Financial Speculation, Crude Oil Futures Volatility, Threshold Cointegration, Granger Causality, Market Instability

**JEL Classification:** C32, Q40, G12
Threshold Cointegration and Dynamics of Crude Oil Futures Volatility and Financial Speculation

Nicholas Ruei-Lin Lee, Hsiang-Hui Chu, Jung-Fang Liu, and Yih-Bey Lin

I. Introduction

World oil demand flat, prices boom. Factors other than supply and demand are now impacting the price. The possible explaining reasons shed light on the growing involvement of speculators in the global oil market. As much as 60% of today’s crude oil price is pure speculation driven by large trader banks and hedge funds as Khan (2009) points out that the ratio of average daily trading volume to global oil production has grown steadily and that, by late 2009, this ratio exceeded 15. In particular, it is common to compare the amount of open

---

1 The US Government’s Energy Information Administration (EIA) reported that, in 2007, China imported 3.2 million barrels per day, and its estimated usage was around 7 million b/d total. The US, by contrast, consumes around 20.7 million b/d. That means the key oil consuming nation, the USA, is experiencing a significant drop in demand. China, which consumes only a third of the oil the US does, will see a minor rise in import demand compared with the total daily world oil output of some 84 million barrels, less than half of a percent of the total demand.

2 For huge US or EU pension funds or banks desperate to get profits following the collapse in earnings since August 2007 and the US real estate crisis, oil is one of the best ways to get huge speculative gains.
interest in the futures market to U.S. daily consumption and use the fact that open interest is many times larger than U.S. daily consumption to infer the presence of speculative pressures (Cho, 2008; Masters, 2008). Therefore, by purchasing large numbers of futures contracts, and thereby pushing up futures prices to even higher levels than current prices, speculators have provided a financial incentive for oil companies to buy even more oil and place it in storage. Crude oil futures volatilities and speculation are of overriding concern.

Traditionally, the Commodity Futures Trading Commission (CFTC) makes a distinction between commercial and non-commercial firms, with commercials being considered “hedgers” and non-commercials “speculators.” Making use of currently available weekly data from CFTC surveys, it is difficult to prove that there exists a significant and daily Granger causality relationship between the rise in the oil price and increasing speculation (Chilton, 2008). On the other hand, but classifying firms in this way is not completely clear cut, because commercials can take speculative positions and non-commercials can hedge exposure to price risk. It is more realistic to think of behaviors of hedgers and speculators as arrayed along a continuum, with some speculating, some hedging and some doing both. Thus, this paper examines speculation taken by both hedgers and speculators in crude oil futures markets, as differed from prior studies.

Understanding the relationship between speculation and volatility for crude oil futures is an important issue. However, few studies focus on the dynamics of crude oil futures price volatility and speculation although prior studies find mixed impacts of financial futures trading

---

3 Chilton (2008) shows that large numbers of non-commercial traders with the purpose of speculative trading are classified as commercial traders by the CFTC. In August, 2008, the CFTC survey showed that an oil speculator, who had been classified as a commercial trader by CFTC was holding 320 million crude oil futures positions, was subsequently reclassified as a non-commercial trader by the CFTC. This made the proportion of non-commercial trader positions increase from nearly 38% to 48%. Consequently, the CFTC revised not only its latest reports, but also all market data reports up to 2007.
on the volatility. Herbert (1995) finds that volume may explain the volatility although Nainar (1993) shows that weekly volatility tends to increase with petroleum futures trading. Herbert also discovers that past levels of volume influence current price volatility but that past variability has much less of an influence on current levels of trading. By contrast, in the study of Foster (1995), there is contemporaneous positive relation between volume and volatility for crude oil futures. Liew and Brooks (1998) find strong evidence of daily, monthly, yearly, volume and open interest effects in volatility in the Kuala Lumpur crude palm oil futures market.

Fleming and Ostdiek (1999) document a strong inverse relation between the open interest and volatility. As splitting volume and open interest into expected and unexpected components, Fleming and Ostdiek (1999) find that, when open interest is greater, the volatility shock associated with a given unexpected increase in volume is much smaller. Fung and Hsieh (2000) show that, if a large number of non-commercial firms (such as hedge funds) use positive feedback trading rules, they can artificially amplify a market move. Ripple and Moosa (2009) show that trading volume and open interest are significant determinants of volatility and support earlier findings of a positive and significant role for trading volume, and they also show the importance of open interest in determining volatility, exerting a significant negative effect. Chevillon and Rifflart (2009), Cifarelli and Paladino (2010), and Kaufmann and Ullman (2009) support that increasing speculation has been one of the important driving forces in the surge of oil prices since 2000. Lee (2010) shows higher speculative and hedging activities in the petroleum futures markets, like crude oil, heating oil, and unleaded gasoline, at the New York Mercantile Exchange (NYMEX) since 2000s but less speculative and hedging behaviors prior to 2000s.

Prior studies examine the lead-lag relation between financial futures trading and price volatility and mix results are found (Shalen, 1993; Fung and Patterson, 1999; Chang et al., 2000; Kocagil and Shachmurove, 1998; Kyriacou and Sarno, 1999; Hsueh et al., 2008). Some studies argue futures trading leading price volatility while others argue the opposite direction, bidirectional relations or no evidence between two variables.
However, little study examines the linearity or threshold-type non-linearities between speculation and volatility in crude oil futures markets although prior studies only examine the linearities between crude oil futures and spot prices using traditional cointegration analysis (Natanelov et al., 2011; Huang et al., 2009; Zhu et al., 2011). A drawback of the traditional cointegration analysis is that this literature fails to account for possible structural breaks in the cointegrating vector, though clearly there is record of structural breaks in crude oil price data, like the increase in the price of crude oil since 2003, its sharp spike at $142 per barrel in July 2008 and its subsequent collapse in the autumn of 2008. Balke and Fornby (1997) first allow the combination of non-linearity and traditional cointegration, but they don’t allow for non-linear adjustment to the long-run equilibrium. The recently developed extension of the above models is proposed by Hansen and Seo (2002). This approach tests for two-regime threshold cointegration and it considers a vector error-correction model (VECM) with one cointegrating vector and a threshold effect based on the error-correction term. Chen et al. (2005) further provide supportive evidence for asymmetric adjustment in weekly U.S. retail gasoline prices. Dunis et al. (2006) also apply threshold cointegration methods to daily closing prices for NYMEX futures contract and unleaded gasoline and find asymmetric adjustment. Similar cointegration methods with daily data are further applied in prior studies of Coleman (1990), Bekiros and Diks (2008), Kellard et al. (2010), and Kühl (2010).

---

*This approach is adopted in several recent papers, revealing threshold-type non-linearities in time series of macroeconomic, for instance, in real exchange rates (Michael et al., 1997; O’Connell, 1998; Aslanidis and Kouretas, 2005; Nakagawa, 2010), in the term structure (Hansen and Seo, 2002), in purchasing power parity doctrine and the law of one price (Enders and Falk, 1998; Baum et al., 2001; Lo and Zivot, 2001), in covered interest parity (Balke and Wohar, 1998) as well as in modeling interest rate policy (Baum and Karasulu, 1997).*
Light sweet crude oil (WTI) futures at the NYMEX are the world's most actively traded energy product.\(^6\) WTI plays an important role in managing risk in the energy sector worldwide because of the most liquidity, most customers, and most transparency. It suggests the expansion of trade activity over these years since the creation of the NYMEX crude oil contract, especially as related to the time structure and the interrelated changes in the absolute prices of oil. Additionally, price trends, however, have been most volatile recently, especially through the 2005 to 2009 time frame, with never before experienced price range movements and record-setting highs before a major collapse in early 2009. In this paper, we examine crude oil futures contracts at the NYMEX and a causal connection between speculation and crude oil futures return volatility.

In this paper we investigate threshold cointegration and dynamics of volatility and speculation for crude oil futures at the NYMEX in the framework of a threshold vector error correction model (TVECM). Two regimes, the usual and unusual regime, are determined by the model. We consider two volatilities measures (VOL), absolute return and high-low volatilities (ABSRV and HLV), and we use the ratio of volume (V) to open interest (OI) as speculation variable (SPEC).

Investigation of the relationship between speculation and crude oil futures return volatility with daily data is important on several reasons. First, the analysis in this paper presents stylized facts about threshold dynamics between crude oil futures return volatility and speculation. Given for a TVECM with daily data, we further understand daily different asymmetric relations in the different two regimes and daily threshold dynamics between volatility and speculation in crude oil futures markets. Second, our findings may also have important policy and trading implications. For example, finding a significant threshold causal relationship between crude oil futures return volatility and speculation may be of interest to market

\(^6\) The NYMEX reports that the total open interest has gone from less than 100,000 contracts in 1986 to peaks over 2009 of almost 1.6 million contracts, reflecting extensive participation from commodity users around the world.
regulators as they decide on the effectiveness or the appropriateness of market restrictions such as daily price movement limits and position limits. However, we expect the existence of threshold effect of crude futures return volatility and speculation during our daily sample period. Crude futures return volatility leading speculation stabilizes the market while reverse relation destabilizes the market. Of the two regimes, the former seems more likely than the latter. It implies that the latter leads to destabilize market, but the former leads to mitigate the destabilization of the market. Therefore, our findings would shed valuable insights on financial implications for threshold dynamics between volatility and speculation.

Our results show the existence of a threshold effects for volatilities (VOL) based on absolute return volatility (ABSRV) and high-low volatility (HLV) and speculation (SPEC) in the crude futures markets by applying the TVECM. Additionally, SPEC and ABSRV as well as HLV variables show different error-correction effects and dynamics, with supporting that there are different and strong asymmetries in the speed of adjustment to the long-run equilibrium in the two regimes. In addition of speculation leading volatility based on former measure (VOL), speculation following volatility is found in the usual regime, suggesting relative market efficient. Bidirectional causalities of speculation and different volatilities measures in the unusual regime, suggesting that market instability mitigated in the unusual regime. That is, NYMEX crude oil futures markets seem to work well although the existence of threshold cointegration for volatility and speculation. Therefore, our results highlight the critical importance of using TVECM in empirical studies on threshold cointegration and dynamics of crude oil futures price volatility and speculation.

The organization of the rest of the paper is as follows. The next section reviews methods developed by Hansen and Seo (2002) and explains the data used, and section 3 summarizes the estimated results. Finally, section 4 concludes the article.

II. Methodology and Data

We follow Hansen and Seo (2002) to model a threshold vector error correction model
(TVECM) of order $l+1$ of the crude oil futures price volatility (VOL) and speculation (SPEC). As a motivation for our multivariate nonlinear modeling, Hansen and Seo (2002) examine a two-regime vector error correction model with a single cointegrating vector and a threshold effect in the error-correction term.

### A. Estimation of the threshold parameters

Let $x_t = (VOL, SPEC)$ be a 2 dimensional vector of time series of the crude oil futures price volatility (VOL) and speculation (SPEC) respectively with $t$ observations. It is assumed that there exists a long-run relationship between these two time series with a cointegrating vector $\beta = (\beta_0, \beta_1)'$. The two regime threshold model where the $\gamma$ is the threshold parameter, takes the following form,

$$
\Delta x_t = \begin{cases} 
A_1'X_{t-1}(\beta) + u_t & \text{if } \omega_{t-1}(\beta) \leq \gamma \\
A_2'X_{t-1}(\beta) + u_t & \text{if } \omega_{t-1}(\beta) > \gamma 
\end{cases} 
$$

where $X_{t-1}(\beta) = (1, \omega_{t-1}(\beta), \Delta x_{t-2}, ... \Delta x_{t-k})'$, $\omega_t(\beta) = \Delta x_t, \Delta x_{t-1}, ... \Delta x_{t-k}$, $\omega_{t-1}(\beta)$ denotes the I(0) error-correction term, the $\gamma$ is the threshold parameter, $X_{t-1}(\beta)$ is $k \times 1$ regressor, and $A$ is $k \times 2$ where $k = 2l + 2$. The error $u_t$ is assumed to be a vector martingale difference sequence with finite covariance matrix, $\Sigma = E(u_tu_t')$. In particular, the estimated coefficients of $\omega_{t-1}$ of each regime denote the different adjustment speeds of the series towards equilibrium. This may be rewritten as

$$
\Delta x_t = A'_1X_{t-1}(\beta) \cdot d_{t,1}(\beta, \gamma) + A'_2X_{t-1}(\beta) \cdot d_{t,2}(\beta, \gamma) + u_t
$$

where $d_{t,1}(\beta, \gamma) = 1(\omega_{t-1}(\beta) \leq \gamma)$ and $d_{t,2}(\beta, \gamma) = 1(\omega_{t-1}(\beta) > \gamma)$, and $1(\cdot)$ denotes the indicator function. The coefficient matrices $A_1$ and $A_2$ govern the dynamics in these regimes. Model (2) allows all coefficients (except the cointegrating vector $\beta$) to switch between these two regimes. In many cases, it may make sense to impose greater parsimony on the model, by
only allowing some coefficients to switch between regimes. This is a special case of (2) where constraints are placed on \((A_1, A_2)\). For example, a model of particular interest only lets the coefficients on the constant and the error correction \(\omega_{t-1}\) to switch, constraining the coefficients on the lagged \(\Delta x_{t-j}\) to be constant across regimes.

The threshold effect only has content if \(0 < P(\omega_{t-1} \leq \gamma) < 1\), otherwise the model simplifies to linear cointegration. We impose this constraint by assuming that \(\pi_0 \leq P(\omega_{t-1} \leq \gamma) \leq 1 - \pi_0\) where \(\pi_0\) is a trimming parameter. For the empirical application, we set \(\pi_0 = 0.05\). The parameters \((\beta, A, \Sigma)\) are estimated by maximum likelihood under the assumption that the errors \(u_t\) are iid Gaussian (using the above normalization on \(\beta\)). Let these estimates be denoted \((\hat{\beta}, \hat{A}, \hat{\Sigma})\). Let \(u_t = \Delta x_t - \hat{A}'X_{t-1}(\hat{\beta})\) be the residual vectors. The Gaussian likelihood for model (2) is

\[
L_n(A_1, A_2, \Sigma, \beta, \gamma) = -\frac{n}{2} \log |\Sigma| - \frac{1}{2} \sum_{t=1}^{n} u_t(A_1, A_2, \beta, \gamma)'\Sigma^{-1} u_t(A_1, A_2, \beta, \gamma) \tag{3}
\]

where \(u_t(A_1, A_2, \beta, \gamma) = \Delta x_t - A'_tX_{t-1}(\beta) \cdot d_1(\beta, \gamma) - A'_tX_{t-1}(\beta) \cdot d_2(\beta, \gamma)\).

The MLE \((\hat{A}_1, \hat{A}_2, \hat{\Sigma}, \hat{\beta}, \hat{\gamma})\) are values which maximize \(L_n(A_1, A_2, \Sigma, \beta, \gamma)\).

It is computationally convenient to first concentrate out \((A_1, A_2, \Sigma)\). That is, hold \((\beta, \gamma)\) fixed and compute the constrained MLE for \((A_1, A_2, \Sigma)\) Additionally,

\[
\hat{A}_1(\beta, \gamma) = (\sum_{t=1}^{n} X_{t-1}(\beta)X_{t-1}(\beta)'d_1(\beta, \gamma))'(\sum_{t=1}^{n} X_{t-1}(\beta)\Delta x_t d_1(\beta, \gamma)) \tag{4}
\]

\[
\hat{A}_2(\beta, \gamma) = (\sum_{t=1}^{n} X_{t-1}(\beta)X_{t-1}(\beta)'d_2(\beta, \gamma))'(\sum_{t=1}^{n} X_{t-1}(\beta)\Delta x_t d_2(\beta, \gamma)) \tag{5}
\]

Note that \(\hat{A}_1(\beta, \gamma)\) and \(\hat{A}_2(\beta, \gamma)\) are the OLS regressions of \(\Delta x_t\) on \(X_{t-1}(\beta)\) for the subsamples for which \(\omega_{t-1} \leq \gamma_{t-1}\) and \(\omega_{t-1} > \gamma\), respectively.
\[
\hat{u}_t(\beta, \gamma) = u_t(\hat{A}_1(\beta, \gamma), \hat{A}_2(\beta, \gamma), \hat{A}_3(\beta, \gamma))
\]

(6)

\[
\hat{\Sigma}(\beta, \gamma) = \frac{1}{n} \sum_{t=1}^{n} \hat{u}_t(\beta, \gamma)\hat{u}_t(\beta, \gamma)'
\]

(7)

This yields the concentrated likelihood function as follows,

\[
L_n(\beta, \gamma) = L_n(\hat{A}_1(\beta, \gamma), \hat{A}_2(\beta, \gamma), \hat{\Sigma}(\beta, \gamma), \beta, \gamma) = -\frac{n}{2} \log |\hat{\Sigma}(\beta, \gamma)| - \frac{np}{2}
\]

(8)

where the MLE \((\hat{\beta}, \hat{\gamma})\) are thus found as the minimizers of \(\log |\hat{\Sigma}(\beta, \gamma)|\) subject to the normalization imposed on \(\beta\) as discussed in aforementioned and previous paragraphs and the constraint \(\pi_0 \leq n^{-1} \sum_{t=1}^{n} l(x_t', \beta, \gamma) \leq 1 - \pi_0\). The MLE for \(A_1\) and \(A_2\) are \(\hat{A}_1 = \hat{A}_1(\beta, \gamma)\) and \(\hat{A}_2 = \hat{A}_2(\beta, \gamma)\).

This criterion function (8) is not smooth, so conventional gradient hill-climbing algorithms are not suitable for its maximization. In the leading case \(p = 2\), Hansen and Seo (2002) suggest using a grid search over the two-dimensional space \((\beta, \gamma)\). In higher dimensional cases, grid search becomes less attractive and alternative search methods might be more appropriate. Note that in the event that \(\beta\) is known a priori, this grid search is greatly simplified. To execute a grid search, one needs to pick a region over the threshold parameter \((\gamma)\) and cointegrating vector \((\beta)\) to joint grid search.

**B. Tests for threshold effects**

Pippenger and Goering (2000) present simulation evidence that linear cointegration tests can have low power to detect threshold cointegration. However, Hansen and Seo (2002) present a quasi-MLE algorithm for constructing estimates of a two-regime threshold
cointegration model and a sup LM statistic for the null hypothesis of no threshold. In addition, they derive an asymptotic null distribution for this statistic and develop methods to calculate the asymptotic distribution by simulation, and how to calculate a bootstrap approximation.

A test for the null of no cointegration in the context of the threshold cointegration model is conducted. We follow Hansen and Seo (2002), and the LM statistic with testing for the presence of the threshold cointegration is employed as follow.

\[
LM(\beta, \gamma) = \text{vec}(\hat{A}_1(\beta, \gamma) - \hat{A}_2(\beta, \gamma))' \times (\hat{V}_1(\beta, \gamma) + \hat{V}_2(\beta, \gamma))^{-1} \times \text{vec}(\hat{A}_1(\beta, \gamma) - \hat{A}_2(\beta, \gamma))
\]

where \( \hat{V}_1(\beta, \gamma) \) and \( \hat{V}_2(\beta, \gamma) \) are defined as the Eicker–White covariance matrix estimators.

To calculate asymptotic critical values with the respective p-values, Hansen and Seo (2002) develop two tests the sup \( LM^0 \) and the sup \( LM \) tests for a given or estimated \( \beta \) using a parametric bootstrap method. The first test is denoted as sup \( LM^0 = \sup_{\gamma_L \leq \gamma \leq \gamma_U} LM(\beta^0, \gamma) \) and would be used when the true cointegrating vector \( \beta \) is known a priori. The second test is used when the true cointegrating vector \( \beta \) is unknown and they denote this test statistic as sup \( LM = \sup_{\gamma_L \leq \gamma \leq \gamma_U} LM(\hat{\beta}, \gamma) \) where \( \beta \) is the null estimate of the cointegrating vector. In these tests, the search region \([\gamma_L, \gamma_U]\) is set so that \( \gamma_L \) is the \( \pi_0 \) percentile of \( \hat{\omega}_{1,1} \) and \( \gamma_U \) is the \((1 - \pi_0) \) percentile.

Finally, the asymptotic distribution depends on the covariance structure of the data, precluding tabulation. Hansen and Seo (2002) suggest using either the fixed regressor bootstrap of Hansen (1996, 2000), or alternatively a parametric residual bootstrap algorithm, to approximate the sampling distribution where the tests were done using a parametric bootstrap method with 1,000 replications.
C. Granger causality test

To conduct Granger causality test for short run dynamics of price volatility (VOL) and speculation (SPEC) in the crude oil futures markets in the TVECM, threshold model with two regimes in the eq. (1) is rewritten as follow.

\[ \Delta VOL_t = \lambda_1 \omega_{i-1} + \alpha_{10} + \sum_{i=1}^{m} \alpha_i \Delta VOL_{t-1} + \sum_{j=1}^{n} \beta_i \Delta SPEC_{t-j} + u_t, \quad \omega_{i-1} \leq \gamma \]  
\[ \Delta SPEC_t = \lambda_2 \omega_{i-1} + \alpha_{20} + \sum_{j=1}^{m} \alpha_j \Delta VOL_{t-j} + \sum_{j=1}^{n} \beta_j \Delta SPEC_{t-j} + u_{2t}, \quad \omega_{i-1} > \gamma \]  

where \( \omega_{i-1} = VOL_{t-1} - \beta \times SPEC_{t-1} \).

Formally, in our TVECM model with two regimes, for short run dynamics of VOL and SPEC, a time series VOL causing another time series SPEC in the Granger sense (denoted as \( VOL \rightarrow SPEC \)) is that even with information about past values of SPEC, one can improve the prediction of SPEC using past values of VOL. If the reverse is true, then we say SPEC Granger-causes VOL (or \( SPEC \rightarrow VOL \)). When both relationships are true, a feedback effect is said to exist between VOL and SPEC.

Finally, we further provide financial implication of the threshold value of \( \gamma \). We compare the threshold value of \( \gamma \) and \( \omega_{i-1} \), where \( \omega_{i-1} = VOL_{t-1} - \beta \times SPEC_{t-1} \). In the usual regime (\( \omega_{i-1} \leq \gamma \)), we expect larger adjustment speed of SPEC or smaller adjustment speed of VOL towards equilibrium but reverse relations in the unusual regime (\( \omega_{i-1} > \gamma \)). It implies that price volatility leading speculation in the usual regime suggests relatively more liquid and efficient market while mixed evidences of price volatility following speculation/bi-directional feedbacks in the unusual regime suggest the destabilized market, even mitigating the destabilization of the market.
D. Data and variables defined

This study employs daily data spanning the period from January 1, 2001 through May 30, 2014. Futures trading volume and open interest across all outstanding contracts, as well as the daily closing prices of the nearby futures contracts on crude oil at the NYMEX are obtained from DataStream database. So the time series of 3,362 sampling points are obtained.

In this study, we follow Garcia et al. (1986) and define speculation (SPEC) as the ratio of daily closing volumes over open interest as follows:

\[ SPEC_t = \frac{V_t}{OI_t} \times 100 \]  

(11)

where \( V_t \) and \( OI_t \) represent aggregate volume traded and open interest of crude oil futures on day \( t \) across all outstanding contracts, respectively. The ratio of volume to open interest captures speculation under the assumption that the majority of speculators and the minority of hedgers prefer to get in and out of the market in a short period of time, in contrast to futures traders who are not engaging in speculation. SPEC measure is to avoid the potential expiration effects on trading activities (Garcia et al., 1986; Luu and Martens, 2003; Hsueh et al., 2008). This higher ratio suggests higher speculative futures activities while the lower one suggests lower speculative futures activities.

Additionally, this measure is used by Luu and Martens (2003) and Hsueh et al. (2008), and they point out that using this ratio to measure speculating trading activities can better capture the information arrival than trading volume or open interest alone. Therefore, changes in this ratio would potentially capture changes in speculative activity.

Next, we use absolute value of crude futures returns (ABSRV) and high low price range (HLV) as return volatilities, as followed.

\[ ABSRV_t = |\ln(P_t / P_{t-1})| \times 100 \]  

(12)

\[ HLV_t = |\ln(H_t / L_t)| \times 100 \]  

(13)

(14)
where $P_{t-1}$ and $P_t$ are daily futures price on two successive days $(t-1)$ and $t$ respectively. $H_t$ and $L_t$ are daily futures high and low price on the day $t$. Variations of this first measure are widely employed in the finance literature (Karpoff, 1987; Garcia et al., 1986; Holthausen et al., 1987; Chan and Chung, 1993; Chatrath and Song, 1998; Gwilym et al., 1999; Ciner, 2002). The second measure of HLV is also applied in prior studies (Karpoff, 1987; Garcia et al., 1986; Chan and Chung, 1993; Chatrath and Song, 1998; Ripple and Moosa, 2009). Alternative two VOL measures should be very easy to apply in practice. In practice they would be of great interest because most historical data is quoted with both the high and low in addition to the close. ABSRV is simple computed while HLV is due to richer information content.

III. Results and Discussions

A. Variability in crude oil futures prices, volatility, and speculation

Figure 1 shows a time series plot of price level and speculation as well as volatility for crude oil futures at the NYMEX. There is record of structural breaks in crude oil price data clearly. Crude oil futures prices were fairly constant up to the early 2003 after which time they exhibit an upward trend. Then, we would observe sharp spike at $142 per barrel in July 2008 and subsequent collapse in the autumn of 2008. Lee (2010) shows structural breaks in the crude oil futures market since the somewhat turbulent time of the 2000s.\(^7\)

\(^7\) Except for a brief correction into early 2007, when prices dropped from mid 2006 peaks near $75/Bbl to about $55/Bbl, prices rose sharply through 2008. The economy was relatively weak in 2007, but that did not stop the upward trek from those early year lows in an unexpected and unabated spike toward the highest level ever in the oil markets. By mid 2008, monthly average peaks rose to about $133/Bbl, with daily peaks approaching $150/Bbl.
Additionally, from Figure 1, we find turbulent time of the 2008s and speculation seems to be fairly constant before 2006 but exhibits an upward trend since 2006. Mamatzakis and Remoundos (2011) show that since 2003 there are quite a few picks in net long non-commercial (speculative) positions in the CFTC whilst oil prices follow a steady upward trend. By contrary, Weiner (2009) does not find any support for any effect from speculation in
oil prices according to CFTC (2009) and IMF (2006) studies. Therefore, the components of speculative trading activities are varied across time.

Finally, we observe two series of volatilities and find obvious violent during 2008s. ABSRV and HLV reach high of 20% per day. Therefore, speculation (SPEC) and volatilities (VOL) seem to be correlated.

**B. Summary statistics**

Table 1 reports summary statistics on speculation (SPEC) based on the ratio of volume (V) to open interest (OI) and price volatility (VOL) based on absolute values of price return volatility (ABSRV) and high-low volatility (HLV), with mean, median, maximum, minimum, standard deviation, correlation, and ADF in level and first difference.

From panel A of this Table, SPEC presents the mean of 36.70% per day and the median of 35.43% as well as the standard deviation of 12.01%, ranging from 1.38% to 102% while the mean (median) for ABSRV and HLV presents 1.82% and 3.13% per day (1.40% and 2.70% per day), ranging from 0 to 16.55% and from 0 to 23% respectively. ABSRV and HLV present the standard deviation of 1.71% and 1.86% respectively. Average values for V and OI are 412,099.96 per day and 1,093,272.26 per day respectively, with V ranging from 5,980 to 1,472,088 and with OI ranging from 392,581 to 1,935,176. Our findings are consistent with prior studies. Caballero et al. (2008) show that the source of the oil price surge is large speculative capital flows that moved into the U.S. oil market. Additionally, Du et al. (2011) show that crude oil futures at the NYMEX are actively traded, and investors would take speculative positions on crude oil futures. Additionally, ADF test for SPEC, ABSRV, and HLV in level are unit root for SPEC but weak stationary for both ABSRV and HLV in first difference. ADF test for SPEC, ABSRV, and HLV in first difference are strong stationary. Similar results for V and OI in first difference are found to be stationary.

From panel B, we observe correlation analyses of SPEC, ABSRV, and HLV and find that correlations between speculation (SPEC) and volatilities (VOL) for ABSRV and HLV are significant and positive with 0.2392 and 0.3374. It presents that an increase in volatility
increases with speculation. Moreover, ABSRV is positively correlated to HLV. Our findings are consistent with Foster (1995).

Table 1 Summary statistics on speculation and volatility

<table>
<thead>
<tr>
<th>Panel A: Summary statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>V</td>
</tr>
<tr>
<td>-----------------------------</td>
</tr>
<tr>
<td>Mean</td>
</tr>
<tr>
<td>Median</td>
</tr>
<tr>
<td>Maximum</td>
</tr>
<tr>
<td>Minimum</td>
</tr>
<tr>
<td>Std. Dev.</td>
</tr>
<tr>
<td>ADF in level †</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel B: Correlation analysis of speculation and volatility</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPEC (%)</td>
</tr>
<tr>
<td>SPEC (%)</td>
</tr>
<tr>
<td>ABSRV (%)</td>
</tr>
<tr>
<td>HLV (%)</td>
</tr>
</tbody>
</table>

Note: ** is at 1% significant level; * is at 5% significant level; † we conduct ADF test with no intercept. SPEC is defined as V/OI ×100 where V and OI are futures volume and open interest respectively; ABSRV is defined as absolute return volatility of |ln(P_t / P_{t-1})|×100. HLV is defined as price range, ln(H_t / L_t)×100, where high price (H_t) and low price (L_t) at the day t.
C. Threshold cointegration test

The presence of a threshold was estimated via the application of the Hansen and Seo (2002) sup LM test (when $\beta$ is estimated). The tests were done using a parametric bootstrap method with 1,000 replications, whereas to select the lag length of the TVECM we have used the Schwartz information criterion (SIC), which is minimum at $l = 8$. The results of the threshold cointegration test are reported in Table 2.

We report the result for the estimated cointegrating vector scenario, and the estimated cointegrating coefficients are $\beta = 0.0869$ for ABSRV vs SPEC and $\beta = 0.0869$ for HLV vs SPEC. The estimated threshold values are $\gamma = 1.2292$ for ABSRV vs SPEC and $\gamma = -16.7579$ for HLV vs SPEC and identifies two regimes with statistically different ECM coefficients. The Wald tests for equality of the ECM coefficients for ABSRV vs SPEC and for HLV vs SPEC are significant ($p$-value is less than 1%).

Table 2  Threshold cointegration test

<table>
<thead>
<tr>
<th>Threshold test</th>
<th>ABSRV vs SPEC</th>
<th>HLV vs SPEC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test statistic value</td>
<td>62.6146</td>
<td>56.9517</td>
</tr>
<tr>
<td>Fixed regressor critical value (C.V.)</td>
<td>(0.010) **</td>
<td>(0.010) **</td>
</tr>
<tr>
<td>Residual bootstrap critical value (C.V.)</td>
<td>11.3795</td>
<td>49.8477</td>
</tr>
<tr>
<td>Threshold estimate</td>
<td>1.2292</td>
<td>-16.7579</td>
</tr>
<tr>
<td>Cointegrating estimate</td>
<td>0.0869</td>
<td>0.9628</td>
</tr>
<tr>
<td>Wald Test for Equality of ECM Coef.</td>
<td>18.3056</td>
<td>10.6522</td>
</tr>
<tr>
<td>Percentage of obs. in the regime</td>
<td>94.29%</td>
<td>94.90%</td>
</tr>
</tbody>
</table>

Note: Standard errors in [ ] and $p$-value in ( ). ** is at 1% significant level; * is at 5% significant level.
For ABSRV vs SPEC, the first, or usual regime, occurs when \((\text{ABSRV}_{t-1} - 0.0869 \times \text{SPEC}_{t-1}) \leq 1.2292\) and includes 94.29% of the observations, whereas the second, or unusual regime, includes the remaining 5.71% of observations and is in place when \((\text{ABSRV}_{t-1} - 0.0869 \times \text{SPEC}_{t-1}) > 1.2292\). Additionally, we turn to observe the HLV vs SPEC. The first, or usual regime, occurs when \((\text{HLV}_{t-1} - 0.9628 \times \text{SPEC}_{t-1}) \leq -16.7579\) and includes 94.9% of the observations, whereas the second, or unusual regime, includes the remaining 5.1% of observations and is in place when \((\text{HLV}_{t-1} - 0.9628 \times \text{SPEC}_{t-1}) > -16.7579\).

D. TVECM results

Table 3 reports the estimated coefficients for the TVECM models and Eicker–White standard errors are reported. However, as we have no formal distribution theory for the parameter estimates and standard errors, these should be interpreted somewhat cautiously. Regime1 represents the usual regime whereas regime2 represents the unusual regime.\(^8\)

We first observe the estimated coefficients of \(\omega_{t-1}\) of each regime, denoting the different adjustment speeds of two series towards equilibrium. For TVECM of ABSRV and SPEC, in the usual regime, our findings show that ABSRV variable show significant and minimal error-correction effects \((-0.1526)\) but SPEC variable show significant and maximal error-correction effects \((0.8067)\). However, both variables show maximal dynamics, with in the usual regime the estimated coefficient showing a substantially larger impact. There are also clustering effects on ABSRV and SPEC respectively. On the other hand, when the gap between ABSRV and SPEC is above a critical threshold \(\gamma = 1.2292\), the error-correction effects of both variables in the equation become mixed results. ABSRV variable show significant and maximal error-correction effects \((-0.7532)\) but SPEC variable show insignificant and minimal

\(^8\) We assume the same length of lags in the usual and unusual regimes because of simple estimating in TEVCM. But we first determine the optimal length of lags according to Schwarz information criterion.
error-correction effects (−0.0313). The estimated coefficient shows a substantially larger impact in the unusual regime. There are little clustering effects on ABSRV and SPEC respectively.

For TVECM of HLV and SPEC, in the usual regime, our findings show that HLV variable show insignificant and minimal error-correction effects (0.0045) but SPEC variable show significant and maximal error-correction effects (0.2101). However, both variables show maximal error-correction effects in the unusual regime, with an insignificant HLV of 0.0497 and a significant SPEC of 0.8805. Additionally, there are clustering effects on HLV and SPEC in both two regimes respectively. Moreover, HLV show maximal dynamics, with in both two regimes the estimated coefficient showing a substantially larger impact on SPEC, particular in unusual regime.

We provide some financial implications. In brief, there is larger adjustment speed of ABSRV towards equilibrium in the unusual regime while larger adjustment speed of SPEC towards equilibrium in the usual regime is found. On the other hand, the positive estimated error-correction effects of both variables for HLV and SPEC in both two regimes are found, particular in the unusual regime.

Two possible reasons are provided. In the usual regime the increase in speculation and the subsequent impact on volatility may be more active participation in crude oil futures markets by speculators, especially short-term speculators, who open and close positions in a relatively short period of time. On the other hand, larger adjustment speed of VOL towards equilibrium in the unusual regime may reflect the exit of both speculators and hedgers taking speculative trading, which may have played a role in the recent crisis.

Therefore, crude oil futures markets perform well, especially in the usual regime, even in the unusual regime.
Table 3  Results on TVECM for VOL vs SPEC and HLV vs SPEC

<table>
<thead>
<tr>
<th>Regime1</th>
<th>Regime2</th>
<th>Regime1</th>
<th>Regime2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dep. Var.</td>
<td>ΔVOL_t</td>
<td>ΔSPEC_t</td>
<td>ΔVOL_t</td>
</tr>
<tr>
<td>λ</td>
<td>-0.1526**</td>
<td>0.8067**</td>
<td>-0.7532**</td>
</tr>
<tr>
<td>α_0</td>
<td>-0.2134**</td>
<td>1.2445**</td>
<td>-0.0189</td>
</tr>
<tr>
<td>α_1</td>
<td>-0.8031**</td>
<td>-0.6495**</td>
<td>-0.2688*</td>
</tr>
<tr>
<td>α_2</td>
<td>-0.6943**</td>
<td>-0.9310**</td>
<td>-0.1880</td>
</tr>
<tr>
<td>α_3</td>
<td>-0.5755**</td>
<td>-1.0375**</td>
<td>-0.2245</td>
</tr>
<tr>
<td>α_4</td>
<td>-0.4806**</td>
<td>-1.0186**</td>
<td>-0.0243</td>
</tr>
<tr>
<td>α_5</td>
<td>-0.3954**</td>
<td>-0.9541**</td>
<td>0.2034</td>
</tr>
<tr>
<td>α_6</td>
<td>-0.2315**</td>
<td>-0.7587**</td>
<td>0.1905</td>
</tr>
<tr>
<td>α_7</td>
<td>-0.1453**</td>
<td>-0.3576*</td>
<td>0.1957</td>
</tr>
<tr>
<td>α_8</td>
<td>-0.0796**</td>
<td>-0.0841</td>
<td>0.1053</td>
</tr>
<tr>
<td>β_1</td>
<td>-0.0111*</td>
<td>-0.3759**</td>
<td>-0.0699**</td>
</tr>
<tr>
<td>β_2</td>
<td>-0.0117*</td>
<td>-0.2422**</td>
<td>-0.0371</td>
</tr>
<tr>
<td>β_3</td>
<td>-0.0125*</td>
<td>-0.2017**</td>
<td>0.0342</td>
</tr>
<tr>
<td>β_4</td>
<td>-0.0111*</td>
<td>-0.0469*</td>
<td>0.0103</td>
</tr>
<tr>
<td>β_5</td>
<td>-0.0029</td>
<td>0.0027</td>
<td>-0.0176</td>
</tr>
<tr>
<td>β_6</td>
<td>-0.0013*</td>
<td>0.0004</td>
<td>-0.0231</td>
</tr>
<tr>
<td>β_7</td>
<td>-0.0126**</td>
<td>-0.1254**</td>
<td>-0.0464</td>
</tr>
<tr>
<td>β_8</td>
<td>-0.0131**</td>
<td>-0.1392**</td>
<td>0.0186</td>
</tr>
</tbody>
</table>

Note: The optimal lags are 8 for the pair of ABSRV and SPEC and 6 for the pair of HLV and SPEC based on SIC respectively. Eicker–White covariance matrix estimation method is used. ** is at 1% significant level. * is at 5% significant level.
In Figure 2, we plot the error-correction effect about the estimated regression functions of SPEC, and VOL, as a function of $\omega_{t-1}$, holding the other variables constant. In this figure 2(a), for ABSRV volatility measure, you can see positive error-correction effect of SPEC and negative error-correction effect of VOL on the left size of the threshold, and on the right of the threshold, flat near-zero error-correction effect of SPEC and sharp and negative error-correction effect of VOL. It suggests larger adjustment speed of VOL in the unusual regime and that of SPEC in the usual regime towards equilibrium individually.

In this figure 2(b), for HLV volatility measure, you can see positive error-correction effect of SPEC and flat near-zero error-correction effect of VOL on the left size of the threshold, and on the right of the threshold, sharp and positive error-correction effect of SPEC and positive error-correction effect of VOL. It suggests larger adjustment speed of VOL and that of SPEC in the unusual regime towards equilibrium individually.

Figure 2  SPEC and VOL based on ABSRV and HLV variance response to error correction
Therefore, responses to error-correction effect of SPEC and ABSRV as well as HLV are dramatically different, particular in the unusual regime. A possible reason is that HLV information which investors prefer to observe is used when taking speculative position.

**E. Dynamics in regime1 and regime2**

We also tested the causality of the two time series for VOL based on ABSRV and HLV and SPEC variables using a Granger causality Wald test (Granger, 1969; Granger et al., 2000), which tests the null hypothesis of no causal relationship between the two time series. The results are presented in Table 4. Regime1 represents the usual regime whereas regime2 represents the unusual regime.

From Table 4, there is evidence that ABSRV Granger causes SPEC in the usual regime while the opposite direction in the usual regime. Additionally, in the unusual regime, there are feedbacks for ABSRV and SPEC. On the other hand, HLV does not Granger cause SPEC in the usual regime while bidirectional causalities of HLV and SPEC in the unusual regime. Hence, as expected, VOL based on ABSRV and HLV shows the evidence of market leadership, whereas SPEC adjusts to long-run equilibrium as a consequence.

Table 4  Dynamics of speculation (SPEC) and volatility (VOL) based on ABSRV and HLV in regime1 and regime2

<table>
<thead>
<tr>
<th></th>
<th>ABSRV</th>
<th>HLV</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>regime1</td>
<td>regime2</td>
</tr>
<tr>
<td>$H_0$: SPEC does not cause VOL</td>
<td>3.4957**</td>
<td>1295.0019**</td>
</tr>
<tr>
<td>$H_0$: VOL does not cause SPEC</td>
<td>38.1691**</td>
<td>4330.0567**</td>
</tr>
</tbody>
</table>

Note: F-statistics are reported. ** is at 1% significant level; * is at 5% significant level.

---
9 This approach is also used by Fung and Patterson (1999) and Hsueh et al. (2008).
Bamber et al. (1999) find that differential interpretations explain a significant amount of the trading occurring in a sample where trading volume is higher than the (firm-specific) non-announcement period average. Our findings in the unusual regime are consistent with informed traders acting on their differential interpretations when there is enough liquidity trading to help camouflage their own information-based trades.

**F. Discussions in regime 1 and regime 2**

We further provide financial implication of the threshold value of \( \gamma \). We compare the threshold value of \( \gamma \) and \( \omega_{t-1} \), where \( \omega_{t-1} = \text{VOL}_{t-1} - \beta \times \text{SPEC}_{t-1} \). We consider that alternative two volatilities (VOL) measures are ABSRV and HLV. Figure 3 plots the series of \( \gamma \) and \( \omega_{t-1} \) for two pairs of ABSRV and SPEC as well as HLV and SPEC.

![Figure 3](image_url)  
**Figure 3** The series of \( \gamma \) and \( \omega_{t-1} \) for two pairs of ABSRV and SPEC as well as HLV and SPEC.
From Figure 3, in the usual regime \((\omega_{t-1} \leq \gamma)\), we find larger adjustment speed of SPEC or smaller adjustment speed of VOL based on ABSRV and HLV towards equilibrium but reverse relations in the unusual regime \((\omega_{t-1} > \gamma)\). It implies that volatility leading speculation in the usual regime suggests relatively more liquid and efficient market. On the other hand, mixed evidences of volatility following speculation and bi-directional feedbacks in the unusual regime suggest the destabilized market, even mitigating the destabilization of the market.

In particular, our findings show destabilized market during 2008 global financial crisis periods and stabilized market during non-global financial crisis periods. It suggests turbulent time of the 2008s and speculation seems to exhibit an upward trend since 2008s. In addition, mixed evidences of volatility following speculation and bi-directional feedbacks in the unusual regime are found. Therefore, our results shed valuable insights on the threshold relation between price volatility and speculation.

IV. Conclusions

This paper examines crude oil futures at the New York Mercantile Exchange (NYMEX) and investigates threshold cointegration and dynamics of crude oil futures price volatility (VOL) and speculation (SPEC) in the framework of a threshold vector error correction model (TVECM). Our daily sample period is from January 1, 2001 to May 30, 2014. We use the ratio of volume over open interest as speculation variable while we consider two volatilities measures, absolute return and high-low volatilities (ABSRV and HLV).

Our findings show that while conventional methods fail to detect significant dynamics between crude oil futures price volatility and speculation in the usual and unusual regimes, application of this extended approach reveals the existence of a threshold cointegration and dynamics in the TVECM. Furthermore, our results of the threshold cointegration test identify two regimes with statistically different ECM coefficients.

Our findings present different error-correction effects and threshold dynamics between ABSRV (HLV) and SPEC. Specially, error-correction effects of ABSRV and SPEC suggest
that the estimated coefficient for ABSRV shows a substantially larger impact in the unusual regimes while the estimated coefficient for SPEC shows a substantially larger impact in the usual regime. In addition, error-correction effects of HLV and SPEC are positive in the two regimes. It implies that there are different and strong asymmetries between the two regimes in the speed of adjustment to the long-run equilibrium for crude oil futures price volatility and speculation. Investors determine different response to error-correction effects of SPEC in term of ABSRV and HLV respectively.

Another implication is that we find different threshold dynamics between ABSRV (HLV) and SPEC. As expected, in the usual regime, the evidence of larger effect on volatility leading speculation than reverse effect stabilizes the market, suggesting relatively more liquid and efficient market. During the unusual regime, there are feedbacks for volatility and speculation, suggesting speculation showing dynamic behavior related to crude oil futures volatility. It implies the destabilized market in the unusual regime, even mitigating the destabilization of the market.

Of importance, during the usual and unusual regime, crude oil futures price volatility has a predominant role in the crude oil futures markets. On the other hand, different magnitudes of feedbacks of volatility and speculation in the unusual regime determine the (de)stabilization of the crude oil futures markets based on SPEC and ABSRV (HLV). However, NYMEX crude oil futures markets seem to work well although the existence of threshold cointegration. Therefore, our results should highlight the critical importance of using TVECM in empirical studies on threshold cointegration and dynamics of crude oil futures return volatility and speculation.

(Received 01 November 2013; Accepted 21 October 2014)
Reference


Fung, H. and G. Patterson, 1999, “The Dynamic Relationship of Volatility, Volume, and


Huang, B. N., C. W. Yang, and M. J. Hwang, 2009, “The Dynamics of a Nonlinear Relationship between Crude Oil Spot and Futures Prices: A Multivariate Threshold


原油期貨波動和金融性投機之門檻共整合動態關係*

李瑞琳**、朱香蕙***、劉榮芳****、林益倍*****

摘要

本文以門檻向量誤差修正模型 (TVECM) 探討紐約商品交易所 (NYMEX) 原油期貨波動性和投機之間門檻共整合動態關係。門檻的區間係由模型內生決定，並定義誤差修正項在門檻下（上）之（非）一般區間。我們以原油期貨交易量相對未平倉合約數作爲投機變數。此外，以原油期貨報酬絕對值與高低價差分別作為波動衡量變數。實證結果顯示，在二個門檻區間，原油期貨波動和投機各自朝向長期均衡有不同的調整速度以及顯著不對稱性結果。此外，我們也發現誤差修正項在門檻下之一般區間，除報酬絕對值波動受投機影響外，波動影響投機的單向因果關係，表示市場具有相對效率。而誤差修正項在門檻上之非一般區間，兩者有回饋現象，代表市場不穩定的惡化現象，具有緩和的傾向。整體上，雖然紐約商品交易所 (NYMEX) 的原油期貨市場存在門檻效果，但

* 作者感謝二位匿名審查者所提供的寶貴意見與評論，使本文修正得更臻完善，文中若仍有疏漏之處，由作者自負全責。
** 朝陽科技大學財務金融系助理教授。
*** 國立暨南國際大學財務金融系助理教授。
**** 嶺東科技大學財政系講師。
***** 朝陽科技大學財務金融系教授・本文聯繫作者。電話：(04)2332-3000 分機 7092，Email：yblin@cyut.edu.tw。
DOI：10.3966/054696002014120096001
Threshold Cointegration and Dynamics of Crude Oil Futures Volatility and Financial Speculation

現有制度使原油期貨市場展現大致良好的效率性。

關鍵詞：金融性投機、原油期貨波動、門檻共整合、因果相關、市場不穩定

JEL 分類代號：C32, Q40, G12