The Relationship between Futures Market Speculation and Spot Market Volatility

Xuemei Xiao

A Thesis
in
The Department
of
John Molson School of Business

Presented in Partial Fulfillment of the Requirements
for the Degree of Master of Science (Finance) at
Concordia University
Montreal, Quebec, Canada

January 2018

CONCORDIA UNIVERSITY

School of Graduate Studies

This is to certify that the thesis prepared

By Xuemei Xiao

Entitled: The Relationship between Futures Market Speculation and Spot Market

and submitted in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE IN ADMINISTRATION (FINANCE)

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final examining committee:

Jisun Yu	Chair
Fassil Nebebe	Examiner
Saif Ullah	Examiner
Latha Shanker	Thesis Supervisor(s)

Approved by Chair of Department or Graduate Program Director

Dean of Faculty

Data February 28, 2018

Abstract

The Relationship between Futures Market Speculation and Spot Market Volatility

This thesis investigates the relationship between speculation in futures markets and expected and unexpected volatility in the spot markets for 21 different commodities. I use the index of adequate speculation, INDADSP, and the index of excess speculation, INDEXSP, developed and estimated by Shanker (2017), to capture the degree of speculation required to meet hedging demand, and the degree of speculation in excess of hedging demand, respectively. For comparison, I also use Working's (1960) speculative index T, as a measure of speculation. I estimate the expected volatility (EV) and unexpected volatility (UEV) of the spot market using a GARCH model. The empirical results indicate that the GJR-GARCH model with a Student's t distribution for the error term is the most appropriate model, among the GARCH-family of models, to capture the volatility of 17 of the 21 spot commodity returns. However, the results of feeder cattle indicate the exists of serial correlation of the residuals for all three GARCH model I used, so I drop it and do the further analysis for the rest of 20 commodities and financial contracts. For each commodity, I create time series of matched weekly indices of speculation, expected volatility and unexpected volatility. Next, I investigate the long-run and short-run relationships between volatilities and speculation using an autoregressive distributed lag model. The results indicate that there is a long term relationship between expected and unexpected volatility and the speculative indices, for all commodities, except the Euro, Eurodollar, and U.S. T-bond, and a short term relationship between volatilities and speculation for all commodities. Finally, I apply the Toda-Yamamoto test to investigate the causal relationship between speculation in futures markets and volatility in spot markets. I find that speculation tends to lead expected volatility more than unexpected volatility for the majority of commodities/financial assets. Expected volatility, rather than unexpected volatility, tends to lead speculation for a majority of commodities/financial assets. There is a bidirectional causality between expected volatility and INDADSP, INDEXSP, and T and between unexpected volatility and INDEXSP for several different commodities and financial assets. However, there is no bidirectional causality between unexpected volatility and the speculative indices INDADSP and T for all 20 commodities/financial assets.

Acknowledgements

First of all, I am very thankful to my supervisor, Professor Latha Shanker, of the Department of Finance, John Molson School of Business, Concordia University. Her suggestions and support during the planning, developing and completion of my thesis were valuable and constructive. I thank Professor Shanker for providing me the use of the speculative indices, the index of adequate speculation and the index of excess speculation that she had developed, estimated and used in her paper, Shanker (2017), which was published in the Journal of Commodity Markets, as well as Working's speculative index T, for the period 1986-2015, in this thesis.

I am grateful to my committee members, Dr. Saif Ullah of the Department of Finance, and Dr. Fassil Nebebe, of the Department of Supply Chains and Business Technology Management, for their useful and constructive comments on this thesis.

Then I am grateful to my parents back in China. They have supported me both spiritually and financially to achieve a Graduate degree, and to experience such a wonderful country and culture.

I am also very thankful to all the friends I have made in this amazing University.

Table of Contents

List of Tables

List of Figures

1. Introduction

Since the opening of a central market place where farmers and dealers could exchange cash for immediate delivery of wheat in Chicago during the 1850s, numerous futures contracts have been developed. Subsequently, farmers (sellers) and dealers (buyers) began to commit to future exchanges of grain for cash to hedge their risk. They also began to close out these agreements before the delivery date. It was not long before people who had no intention of ever buying or selling wheat began trading the contracts-- that was the time when speculators appeared. The key reason for the growth in futures markets is uncertainty in spot prices. Meanwhile, futures markets also allow people with different beliefs to speculate on those beliefs and thereby affect spot and futures prices.

Futures trading has boomed in terms of the volume and the number of participants since the 1970s. During the first decade of the 21st century, the structural changes in commodity futures markets were more than ever before (Irwin and Sanders, 2012b). The available literature indicates that the role of speculators in derivatives markets had created a "bubble" in commodity prices, which greatly exceeded their fundamental values. When the "bubble" burst, the results were catastrophic. Irwin and Sanders (2012) referred to the rapid growth in held positions and traded volume as the "financialization" of commodity markets. The public opinion that was held in many countries was that futures prices were destabilized by speculative and hedging pressure from futures market participants. However, Friedman's (1953) theory was that profitable speculation must involve buying when the price is low and selling when the price is high so that irrational speculators or noise traders trading on irrelevant information will not survive in the market place. Friedman's viewpoint was that speculation would necessarily stabilize markets. Therefore, the question of whether futures market speculation destabilizes or stabilizes spot markets is an empirical issue.

The motivation for this thesis are as follows. Volatility plays a significant role in the pricing of options and futures and it is important for traders to pay close attention to a market's volatility as they build their strategies. Most studies of the relationship between futures market speculation and spot market volatility focus on the short time period. Hence, I propose to address the following two questions: 1) What is the most appropriate model to estimate the spot price volatility for 21 commodities and financial assets? 2) What is the relationship between speculation in futures markets and expected and unexpected volatility in the underlying spot markets, both in the short term and in the long term? 3) What is the causal relationship between speculation in the futures markets and volatility in the underlying spot markets?

I use the index of adequate speculation INDADSP and the index of excess speculation INDEXSP, developed and estimated by Shanker (2017) and Working's (1960) speculative index T, to capture speculative activity in 21 different futures contracts for the period 1986-2015[1]. I model the expected and unexpected volatilities for the corresponding spot returns for the same period as a GARCH process. I model the long-run and short-run relationships between volatilities and speculation as an autoregressive distributed lag (ARDL) model. Finally, I apply the Toda-Yamamoto causality tests on expected and unexpected spot return volatilities and futures market speculation to determine the causal relationship, if there is any, between them.

2. Literature review:

Whether futures market speculation destabilized or stabilized spot markets has been addressed both theoretically as well as empirically. On the theoretical side, Cox (1979) and Chatrath et al. (1996) point out that because of lower transaction costs and higher leverage in futures market as compared to spot markets, uninformed speculative investors tend to shift from the spot market to the futures market. This trend will decrease market depth in spot markets and increase their volatility. However, other studies have pointed out that futures market speculation stabilizes the spot market (Danthine, 1978; Kyle, 1985, and Froot and Perold, 1995). In their studies, they argued that futures markets provide market depth, a platform for hedging and arbitrage and facilitate price discovery, which improves market efficiency and thus stabilizes spot markets. On the empirical side, researchers focus on two aspects, which are: 1) investigating spot market volatility before and after the introduction of futures markets (Antoniou et al, 1998; Lee and Ohk, 1992) and: 2) examining the interaction between futures trading activity and spot market volatility. (Bessemhinder and Seguin, 1992 and Gulen and Mayhew, 2000).

When studying the interaction between futures trading activity and spot market volatility, researchers examined the relationships in different categories of futures markets and different countries. The results varied according to these differences. In equity markets, Bessembinder and Seguin (1992) decomposed the trading volume and open interest into expected and unexpected components and found that stock market volatility is positively related to unexpected trading activity in futures markets, but negatively to expected trading activity. On the other hand, Chang et al. (2000) divided volatility estimates into expected and unexpected components to investigate whether traders' reactions to volatility depend upon its predictability. They found that hedging activity in futures markets has a positive relationship with unexpected volatility. However, there is no relationship between speculative activity and futures market volatility and when there is a higher volatility in the stock market, the participation of hedgers is significantly larger than that of

[1] Data on the indices of adequate and excess speculation and the T index, for the period 1986-2015, were provided by Professor Shanker, for my use in this thesis.

speculators. In currency markets, Clifton (1985), Chatrath (1996), Grammatikos and Saunders (1986), and McCarthy and Najand (1993) found a positive correlation between spot price volatility and the volume of futures trading. However, Adrangi and Chatrath (1998) and Sarwar (2003) found a stabilizing effect of futures trading on currency markets. In commodity markets, Pashigian (1986) and Weaver and Banerjee (1990) found that futures trading activity destabilizes the spot return volatility of agricultural commodities. Yang et al. (2005) examined the dynamic lead-lag relationships between futures trading volume, open interest and spot volatility for agricultural commodities and found that an unexpected increase in futures trading volume unidirectionally causes an increase in spot price volatility for most agricultural commodities. However, they found a weak causal relationship between open interest and spot volatility. Lehecka (2013) investigated the relationships between trading positions and prices in commodity futures market using Toda-Yamamoto causality tests for 24 commodities from 2006 to 2011. He concludes that there is little evidence of a systematic lead-lag relationship from hedging and speculative activity to prices in commodity markets. However the results indicate that prices in commodity markets may cause traders to change their positions.

2.1 Measures of Futures Market Activity:

One way to measure futures market activity is through trading volume and open interest. Trading volume in futures market is the number of futures contracts being exchanged between buyers and sellers. Open interest is the number of futures contract that are open and held by traders and investors. It is a measure of the flow of money into or out of a futures market. Increasing open interest represents new or additional money coming into the market. Bessembinder and Seguin (1992, 1993) investigated stock index futures markets by decomposing the trading volume and open interest into expected and unexpected components using an Autoregressive Integrated Moving Average (ARIMA (p, q)) model. They defined the fitted value from the ARIMA model as the expected components (of volume and open interest) and the residuals as unexpected components. Then they used expected and unexpected volume and open interest as four explanatory and exogenous variables in the spot volatility equation through an augmented GARCH process. This model has been widely used in the finance literature in papers such as Chatrath et al. (2003), Gulen and Mayhew (2000) and Lee, Stevenson and Lee (2014).

However, in the study of aggregate demand in futures markets, open interest is a more appropriate measure of participant activity than trading volume, since trading volume is more transitory in nature (Chang et al., 2000). There have been broad studies which attribute the sharp change in futures prices to hedging and speculative pressure. The concept of hedging pressure is introduced based on the theory of “normal backwardation" of Keynes (1930) and Hicks (1939). It assumes that if the demand for short hedging is more than the demand for long hedging, this gap needs to be fulfilled by an additional return risk premium to the long speculators, which may influence prices. Speculative pressure occurs when the demand for long speculation exceeds short hedging needs. A number of studies showed that both hedging pressure (e.g., Bessembinder 1992; De Roon et al. 2000 and Basu and Miffre 2009) and speculative pressure (e.g. Gilbert 2010a 2010b and Singleton 2011) may change the futures prices. However, some studies (e.g., Wang 2003; Bryant et al. 2006 and Sanders et al. 2009) did not find any evidence that either hedgers' or speculators' positions lead futures prices. There are a number of studies that found evidence of contemporaneous relationships between the positions of market participants and futures prices, but the causal linkages were weak.

One of the reasons why these researchers obtained different results on the relationship between trading activities and price volatility is that they used different measures of hedging and speculative activity. In order to measure trading activity in futures markets, basically, we need to know the futures positions of market participants. All futures positions of market participants are publicly available from the Commodity Futures Trading Commission (CFTC)'s reports in its Commitments of Traders (COT) reports, on a weekly basis. Every Tuesday, the futures market's open interest positions of market participants are collected by aggregating across all contract expiration months for a given commodity. The following Friday at 3:30 p.m. EST, the data is made available to the public. (Lehecka, 2013).

According to CFTC practice, traders in futures markets are classified as reporting and non-reporting based on the size of their positions. Reporting traders who dominate the open interest of futures markets (70% - 90% of the open interest of any futures markets) and hold positions in excess of the CFTC reporting level, are further classified as commercial (hedgers) or non-commercial (speculators) traders. Therefore, the market's total open interest (TOI) is disaggregated in the following way:

$$\underbrace{\overbrace{[\mathrm{NCL} + \mathrm{NCS} + 2 * \mathrm{NCSP}]}^{\text{Noncommercial}} + \overbrace{[\mathrm{CL} + \mathrm{CS}]}^{\text{Commercial}}}_{\text{Reporting}} + \underbrace{[\mathrm{NRL} + \mathrm{NRS}]}_{\text{Nonreporting}} = 2 * \mathrm{TOI}, \tag{2.1}$$

where NCL, NCS and NCSP are non-commercial long, short, and spreading positions, respectively. CL (CS) represents commercial long (short) positions, and NRL (NRS) are long (short) positions held by non-reporting traders.

In the paper of De Roon et al. (2000) and Basu and Mire (2009), the "hedging pressure (HP_t)" variable, which captures the net long position of commercials as a percentage of the total position of commercials is defined as:

$$\mathrm{HP_t} = \frac{\mathrm{CL_t} - \mathrm{CS_t}}{\mathrm{CL_t} + \mathrm{CS_t}}. \tag{2.2}$$

Following De Roon et al. (2000), the "speculative pressure (SP_t)" variable which captures the net long position held by non-commercials as a percentage of the total position of non-commercial is defined as follows:

$$\mathrm{SP_t} = \frac{\mathrm{NCL_t} - \mathrm{NCS_t}}{\mathrm{NCL_t} + \mathrm{NCS_t} + 2 * \mathrm{NCSP_t}}, \tag{2.3}$$

and the "small trader pressure (STP_t)" variable which captures the net long position of non-reporting traders as a percentage of the total position of non-reporting traders is defined as:

$$\mathrm{STP_t} = \frac{\mathrm{NRL_t} - \mathrm{NRS_t}}{\mathrm{NRL_t} + \mathrm{NRS_t}}. \tag{2.4}$$

Hedging, speculative and small trader pressure variables are bound to be between -1 and 1. For example, a value for hedging pressure of 0.4 indicates that over the previous week 40% of hedging positions were net long, while, a value for speculator pressure of -0.4 indicates that over the previous week 40% of speculative positions were net short. $\mathrm{HP_t}$, $\mathrm{SP_t}$, and $\mathrm{STP_t}$, weighted by their present of TOI, will sum to zero.

Working (1960) developed a speculative index T to measure the intensity of speculation relative to hedging. Since long and short hedgers will not always trade at the same time or in the same quantity, speculators are needed to meet hedging demand. The T index is defined as the ratio of short speculation (SS) (or long speculation (SL) depending on the relative volumes of short hedging and long hedging) to the sum of short hedging (HS) and long hedging (HL) positions. The speculative index T is constructed as following:

$$T\begin{cases}1+\dfrac{SS}{HS+HL} \ if \ HS>HL \\ 1+\dfrac{SL}{HS+HL} \ if \ HL>HS\end{cases}. \tag{2.5}$$

The T index has a minimum value of 1.00, when the level of speculation equals hedging needs. According to Working, a value for the T index of 1.20 indicates that speculation is 20% in excess of what is necessary to meet hedging needs.

Shanker (2017) offered two new indices, an index of adequate speculation, *INDADSP*, which measures the degree of speculation needed to meet unbalanced hedging, and an index of excess speculation, *INDEXSP*, which measures the degree of speculation in excess of that required to meet unbalanced hedging. Shanker's (2017) formulas for the indices are provided in **Eq (2.6)** and **Eq (2.7)** which follow:

$$\begin{cases}INDADSP = 1-\left(\dfrac{HB}{HL}\right)*\left(\dfrac{HL}{HS}\right), & if \ HS \geq HL \\ INDADSP = 1-\left(\dfrac{HB}{HS}\right)*\left(\dfrac{HS}{HL}\right), & if \ HS \leq HL\end{cases} \tag{2.6}$$

$$\begin{cases}INDEXSP = \dfrac{SB}{HS} = \dfrac{SL}{HS} - INDADSP, & if \ HS \geq HL \\ INDEXSP = \dfrac{SB}{HL} = \dfrac{SL}{HL} - INDADSP, & if \ HS \leq HL\end{cases} \tag{2.7}$$

In **Eq (2.6)** and **Eq (2.7)**, *HS, HL, HB, SS, SL* and *SB* represent open short hedging, long hedging, balancing hedging contracts, short speculation, long speculation and balancing speculative contracts, respectively. For example, INDEXSP = 0.2 means that speculation is 20% in excess of that required to meet unbalanced hedging.

Shanker (2017) noted that these two indices together correctly estimate Working's conceptual definition for a speculative index, which is the ratio of speculation to unbalanced hedging. She noted the shortcomings of Working's formula for the speculative index T. She pointed out that "Working's formula for the speculative index T is difficult to explain for markets with long hedging, does not explicitly incorporate balancing hedging, accurately measures his conceptual definition only for a market with no long hedging, and implies that excess speculation exists in markets in which it is absent". Working recognized that the formula for T will equal 1 if only there are no short speculators in the market, implying that the futures price would be “so low that no speculator thought it would go lower”, which Working judges is “too low”. Working concluded, in this case, that more speculation is “economically necessary” than required to meet unbalanced hedging. However, Shanker noted that this only follows because of the error in Working’s formula for the speculative index T, which is provided in **Eq (2.5)**. When Working’s conceptual definition for the speculative index is correctly estimated, using the indices of adequate and excess speculation, the implication is that long speculation could equal unbalanced short hedging and short speculation would still exist in the market, and would equal unbalanced long hedging.

Shanker (2017) estimated the indices of adequate and excess speculation for 21 different futures contracts for the period 1986-2015. She used the indices of adequate and excess speculation to investigate the relationship between speculation and volatility of the NYMEX’s West Texas Intermediate (WTI) crude oil futures contract over the period 31 January 1986 till 29 December 2015, while controlling for market fundamental risk. In order to estimate the unobservable variables -- balancing hedging and balancing speculative contracts, which are necessary to estimate the indices, she applied a Kalman (1960) filter approach with inequality constraints imposed on the state variables, which are the time-varying intercept and slope of the true linear relationship between the speculative and hedging ratio for each contract.

Wang (2001) also developed a sentiment index based on COT positions in six actively traded agricultural futures markets. His sentiment index, $SI_{i,t}$ is also calculated for each trader type, i (non-commercial, commercial and non-reporting), based on current aggregate positions and historical extreme values over the previous three years. Wang defined the sentiment index of trader type i in market j at week t as:

$$SI_{i,t}^{j} = \frac{S_{i,t}^{j} - min(S_{i,t}^{j})}{max(S_{i,t}^{j}) - min(S_{i,t}^{j})}, \qquad (2.8)$$

where $S_{i,t}^{j}$ is the aggregate position for trader type i at week t and it is defined as the total long open interest minus the total short open interest. $max(S_{i,t}^{j})$ and $min(S_{i,t}^{j})$ represent historical maximum and minimum aggregate positions for trader type i in market j over the previous three years.

The sentiment index, $SI_{i,t}$ is similar to other investor sentiment indices within the market and it is widely accepted by futures participants. This measure provides a more intuitive analysis of the actions of traders than the number of long or short contracts. Moreover, this index can be used to compare the return predictability across futures markets.

2.2 Modeling Spot Price Volatility:

It is now widely agreed that price volatilities of financial assets are time varying, with persistent dynamics (Poon and Granger, 2005). The common method to capture the time varying nature of volatility is to model the conditional variance as a Generalized Autoregressive Conditional Heteroscedasticity (GARCH) process, first proposed by Engle (1982). A linear regression performed via ordinary least squares (OLS) is considered to be the Best Linear Unbiased Estimator (BLUE) of the regression coefficients. It assumes that the error terms in the regression equation have a constant variance ($Var[\varepsilon_i|X] = h^2, \forall i$) and are uncorrelated with each other ($Cov[\varepsilon_i, \varepsilon_j|X] = 0$). However, when doing empirical studies that utilize time series of returns it is often found that the variance of returns is not constant over time. Over a sample of returns, one can expect to find periods of high volatility and other periods of low volatility. This reduces the effectiveness of the use of OLS in such situations.

In order to address this problem, Engle (1982) suggested that past information which contains the realized value of all relevant variables up to time $t-1$ would affect investors when they made their investment decisions at time $t-1$ as well as the expected return and volatility at time t. Let R_t be the rate of return from time $t-1$ to time t and F_{t-1} represent the past information available up to time $t-1$. Then let m_t and h_t^2 denote the conditional expected value of R_t, given F_{t-1} and the conditional variance of R_t, given F_{t-1} respectively. That is, $m_t \equiv E(R_t|F_{t-1})$ and $h_t^2 \equiv Var(R_t|F_{t-1})$. The unexpected return at time t is $\varepsilon_t \equiv R_t - m_t$. The ARCH (p) model suggests that h_t can be modeled as a function of the lagged ε_ts. Bollerslev (1986) generalizes the ARCH (p) model to the GARCH (p, q) model. Empirically, the family of GARCH models has been very successful. Among these models, the GARCH (1,1) is used most widely. The following is the equation for the GARCH (1, 1) with normal distribution model:

$$R_t = a + \sum_{i=1}^{n} b_i R_{t-i} + \varepsilon_t, (\varepsilon_t|\mathrm{F}_{t-1}) \sim N(0, h_t) \tag{2.9}$$

$$h_t^2 = \omega + \sum_{i=1}^{p} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j\, h_{t-j}^2. \tag{2.10}$$

Although the standard GARCH model has been proven to be successful in modeling the time varying volatility of financial returns, it does not take into account for asymmetry in the distribution of returns, which means that volatility might be higher or lower depending on the sign of previous errors (Black, 1976; French et al., 1987 and Nelson, 1991). For example, bad news ($\varepsilon_{t-1} < 0$) and good news ($\varepsilon_{t-1} \geq 0$) might have different effects on conditional variance. One model to account for this asymmetry is the GJR-GARCH model following the work of Glosten, Jagannathan and Runkle (1993), according to which:

$$h_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1}^2 + \gamma_1 S_{t-1} \varepsilon_{t-1}^2, \tag{2.11}$$

where

$$S_{t-1} = \begin{cases} 1, & if\ \varepsilon_{t-1} < 0 \\ 0, & if\ \varepsilon_{t-1} \geq 0 \end{cases} \tag{2.12}$$

If the leverage effect exists, γ_1 is expected to be positive. Good news has an impact of α_1, while bad news has an impact of $(\alpha_1 + \gamma_1)$. Another GARCH model which can capture the asymmetry between up and down moves is the EGARCH model of Nelson (1991). The asymmetry can be explained by the fact that negative shocks will have a greater impact on volatility than positive shocks. An EGARCH (p, q) model is represented by:

$$\ln(h_t^2) = \alpha_0 + \sum_{i=1}^{q} \alpha_i \frac{\varepsilon_{t-\mathrm{i}} + \gamma_\mathrm{i} |\varepsilon_{t-\mathrm{i}}|}{h_{t-\mathrm{i}}} + \sum_{i=1}^{p} \beta_\mathrm{j} \ln\left(h_{t-\mathrm{j}}^2\right). \tag{2.13}$$

When $\varepsilon_{t-\mathrm{i}}$ is positive or there is good news, the total effect of ε_{t-1} is $(1+\gamma_\mathrm{i})|\varepsilon_{t-\mathrm{i}}|$. However, when $\varepsilon_{t-\mathrm{i}}$ is negative, the total effect of ε_{t-1} is $(1-\gamma_\mathrm{i})|\varepsilon_{t-\mathrm{i}}|$. The null hypothesis of no asymmetric effects can be expressed by $\gamma_\mathrm{i} = 0$. If the hypothesis is rejected, there is asymmetric effect of price shocks. The $\varepsilon_{t-\mathrm{i}}$ mentioned in Eq**(2.12)** and Eq**(2.13)** are both follow normal distribution. However, the returns of financial assets are found to have skewness, excess kurtosis and fat-tails in the distribution. To accommodate these phenomenon, Bollerslev (1987) first combined with Student distribution errors with a GARCH model.

In order to determine which GARCH specification is a better fit to the data, there are several diagnostic checks that could be used. The first one is the Akaike Information Criterion (AIC), which makes adjustments to the likelihood function to account for the number of parameters and estimates the quality of each model for a given set of data, relative to each of the other models. If the number of parameters in the model is P, the AIC is given by:

$$AIC(P) = 2P - 2\ln(maximum\ likelihood). \tag{2.14}$$

Given a set of candidate models for the data, the preferred model is the one with the minimum AIC value. A second criterion is the Schwartz Bayesian Criterion (SBC):

$$SBC(P) = P\ln(N) - 2\ln(maximum\ likelihood), \tag{2.15}$$

where N is the number of observations. It is based, in part, on the likelihood function and it is closely related to the AIC but with a different penalty for the number of parameters. The model with the lowest SBC is preferred.

Another diagnostic check of a model is to compute the residuals of a GARCH specification model and test whether these residuals are i.i.d. The test was proposed by Box and Pierce (1970) and was defined as:

$$Q_{BP} = n\sum_{k=1}^{h}\hat{\rho}_k^2, \tag{2.16}$$

The null hypothesis of existence of autocorrelation are as follows:

$$\begin{aligned} &H_0: \hat{\rho} = 0 \\ &H_1: \hat{\rho} \neq 0. \end{aligned} \tag{2.17}$$

where Q_{BP} is the Box-Pierce test statistic, n is the sample size, $\hat{\rho}_k^2$ is the sample autocorrelation at lag k, and h is the number of lags being tested. Essentially, the Box-Pierce test indicates that if the residuals are white noise, the Q-statistic follows a χ^2distribution with h degree of freedom. For significance level α, the critical region for rejection of the hypothesis of randomness of residuals is:

$$Q_{BP} > \chi^2_{1-\alpha,h}, \tag{2.18}$$

where $\chi^2_{1-\alpha,h}$ is the α-quantile of the chi-squared distribution with h degree of freedom. A modified version was proposed by Ljung and Box (1978), which is represented as:

$$Q = n(n+2)\sum_{k=1}^{h}\frac{\hat{\rho}_k^2}{n-k}. \tag{2.19}$$

It uses the same hypothesis as defined in **Eq (2.17)** and same critical region as defined in **Eq (2.18)**. Simulation studies have shown that the Ljung-Box statistic is better for all sample sizes including small sizes.

By using these diagnostic checks, Gulen and Mayhew (2000) tested the symmetric GARCH (1,1) model, asymmetric GJR-GARCH model, the nonlinear GARCH model (NGARCH; Engle and Ng, 1993), and the EGARCH model to determine which type of GARCH model could be used to model world-wide stock index returns. They found that the GJR-GARCH model performs marginally better than the others. Engle and Ng (1993) also compared the GARCH (1, 1) model with several other volatility models that allow for asymmetry in the impact of news on volatility. They suggested that the GJR-GARCH model was the best model to capture and forecast financial market volatility.

2.3 Cointegration test: Autoregressive-Distributed Lag (ARDL) test:

Before investigating any of the time series, it is necessary to examine the stationarity of each series first. A linear combination of non-stationary time series will lead to spurious regression, since it results in higher t-values and coefficient of determination (R^2) and lower Durbin Watson statistics. All of these will lead to a high frequency of Type I error and biased estimation of the regression coefficients (Granger and Newbold, 1974). Therefore, the Augmented Dickey-Fuller (ADF) test (Dickey and Fuller, 1979) can be used to test for the presence of unit roots. If the variables exhibit different lag orders, the traditional cointegration approaches such as those of Engle and Granger (1987) and Johansen and Juselius (1990) cannot be used. I therefore use the more recent and advanced Autoregressive-Distributed Lag (ARDL) procedure, also referred to as the Bounds test, developed by Pesaran et al. (2001), to find the long run contemporary relationship among the variables. The term "distributed lag" indicates the inclusion of unrestricted lags of the regressors in a regression function. ARDL is preferable when dealing with variables that are integrated of different orders, I(0), I(1) or combination of both. (Pesaran and Pesaran, 1997; Pesaran et al, 2000; Pesaran et al, 2001). Appropriate modifications of the order of the ARDL model can generally provide unbiased and efficient estimates of the long-run model and valid t-statistics even when some of the regressors are endogenous (Pesaran and Shin, 1999; Harris and Sollis, 2003), since the residuals are uncorrelated. The ARDL procedure also estimates the short-term and long-term components of the model simultaneously, removing problems associated with omitted variables and autocorrelation. The ARDL model is widely used in long term macroeconomic analysis. (Alimi, 2014; Ibrahim, 2015)

2.4 Augmented Granger-causality test:

There are numerous methods to test the causal relationship between two variables. Among them, the Granger-Causality proposed by Granger (1969) is the most commonly used method. The concept of this test is captured by: "X is said to Granger-cause Y if Y can be better predicted using the histories of both X and Y than it can by using the history of Y alone." (Granger, 1969) The procedure estimates the following vector autoregressions (VAR):

$$X_t = \sum_{i=1}^{n} \alpha_i Y_{t-i} + \sum_{j=1}^{n} \beta_j X_{t-j} + \mu_{1t} \tag{2.20}$$

$$Y_t = \sum_{i=1}^{m} \theta_i X_{t-i} + \sum_{j=1}^{m} \delta_j Y_{t-j} + \mu_{2t}. \tag{2.21}$$

X_t and Y_t are time series variables at time t, X_{t-i} and Y_{t-j} are the same series variable at $t-i$ and $t-j$ respectively. μ_{yt} and μ_{xt} are error terms that are assumed to be white noise with zero mean, constant variance and no autocorrelation. The assumption under the traditional Granger-causality test is that the disturbances μ_{1t} and μ_{2t} are uncorrelated and both X_t and Y_t are stationary. The null hypotheses of equation **(2.20)** and **(2.21)** is that the estimated lagged coefficient α_i and θ_i are both equal to zero and can be tested with an F-test. If the joint test rejects the two null hypotheses, a causal relationship between X_t and Y_t are confirmed. As noted, the Granger test works well between two variables, however, it may produce spurious results when applied to multiple variables (Maddala, 2001). If any of the variables are non-stationary, whether or not they are cointegrated, the Wald test statistic for this test will not have an asymptotic Chi-Square distribution (Toda and Phillips, 1994). Gujarati (2006) also proved that "when the variables are integrated, the F-test procedure is not valid, as the test statistics do not have a standard distribution."

Although a multivariate VAR model can deal with the first deficiency mentioned above, it still suffers from non-stationarity and can even make it worse, since it is hard to find the same stochastic trend in different time series. Hence Toda and Yamamoto (1995) developed an alternative test, which is applicable whether Y_t and X_t are $I(0), I(1)$ or $I(2)$ are non-cointegrated or cointegrated of an arbitrary order. Their method involves a modified Wald statistic for testing the significance of the parameters of a VAR model which guarantees the asymptotic χ^2 distribution of the Wald statistic. Toda and Yamamoto's (1995) augmented Granger causality test is based on the following equations:

$$Y_t = \alpha + \sum_{i=1}^{h+d_{max}} \beta_i Y_{t-i} + \sum_{j=1}^{k+d_{max}} \gamma_j X_{t-j} + \mu_{yt} \tag{2.22}$$

$$X_t = \alpha + \sum_{i=1}^{h+d} \theta_i X_{t-i} + \sum_{j=1}^{k+d} \delta_j Y_{t-j} + \mu_{xt}, \tag{2.23}$$

where d_{max} is the maximal order of integration of the variables in the system and it can be tested by using an Augmented Dickey-Fuller (ADF) or Phillips-Perron (PP) test. For example, if Y_t and X_t have different integration orders, say $I(1)$ *and* $I(2)$, then, the maximal order of integration is 2. h and k are the optimal lag length of Y_t and X_t respectively, which can be determined by the AIC or the SBC criteria, as defined in **Eq (2.14)** and **Eq (2.15)**. μ_{yt} and μ_{xt} are error terms that are assumed to be white noise with zero mean, constant variance and no autocorrelation. This lag-augmented Granger-causality approach is valid even if the time series variables are difference-stationary. Toda and Yamamoto's (1995) augmented Granger causality test method is widely used in time series analysis (Lehecka, 2013; Dufour et al., 2006; Umar and Dahalan, 2016; Alimi and Ibironke, 2012; Anguibi et al., 2015 and Alimi and Ofonyelu, 2013).

3. Data:

3.1 Spot prices data:

I address the relationship between futures market speculation and spot market volatility for 7 different groups of commodities/financial assets. These groups are: (1) Energy: Crude oil, Nature gas and Heating oil; (2) Agriculture: Soybean, Corn and Wheat; (3) Metal: Gold, Silver and Copper; and (4) Livestock: Feeder cattle, Lean hogs and Live cattle; (5) Foreign exchange: British pounds, Euro and Japanese yen; (6) Fixed-income: Eurodollar, 10-Year T-note and U.S. T-bond; (7) Stock index: DJIA, NASDAQ 100 and S&P 500. The data set consists of daily closing spot prices of the underlying commodities/assets for the period January 1986 to December 2015. Details of the data period and source of data are given in **Table 3.1.1.** All data are provided by Datastream.

Table 3.1.1 Description of the Data

Commodity	Cash Market	Futures Market
Crude oil	West Texas Intermediate Spot Cushing USD / Barrel Source: Thomson Reuters Symbol: CRUDOIL	NYMEX
Natural gas	Henry Hub USD / MMBTU (1990-2015) Source: Thomson Reuters Symbol: NATGHEN	NYMEX
Heating oil	Heating Oil No.2 NYH Spot FPB USD/Gallon (1986-06-06) Source: EIA-Energy Information Administration, United States Symbol: EIAHONY	NYMEX
Corn	No. 2 yellow corn USD / Bushel Source: U.S. Department of Agriculture Symbol: CORNUS2	CBOT
Soybean	No.1 yellow soybeans USD / Bushel Source: U.S. Department of Agriculture Symbol: SOYBEAN	CBOT
Wheat	No.2 Soft Red USD / Bushel Source: U.S. Department of Agriculture Symbol: WHEATSF	CBOT
Feeder cattle	Feeder Cattle Index USD / Points (1993-2015) Source: CME-Chicago Mercantile Exchange Symbol: CFCINDX	CME
Lean hog	S&P GSCI lean hog price Spot index USD / Points Source: CME-Chicago Mercantile Exchange Symbol: CLHINDX	CME
Live cattle	S&P GSCI Live Cattle Spot index USD / Points Source: S&P Symbol: GSLCSPT	CME
Copper	S&P GSCI Copper Spot Index Source: S&P Symbol: GSICSPT	CMX
Gold	Gold, Handy and Harman Base USD / Troy Ounce Source: Handy & Harman Symbol: GOLDHAR	CMX
Silver	Silver, Handy and Harman Base USD / Troy Ounce	CMX

	Source: Handy & Harman Symbol: SILVERH	
DJIA	Dow Jones Industrials Index Source: Dow Jones Symbol: DJINDUS	CBOT
NASDAQ 100	NASDAQ 100 Source: NASDAQ Stock Market Symbol: NASA 100	CME
S&P 500	S&P 500 Composite Source: S&P Symbol: S&PCOMP	CME
Eurodollar	IBA USD Interbank 3M LIBOR Source: ICE Benchmark Administration Ltd. Symbol: BBUSD3M	CME
10-Year T-note	US Government Benchmark Constant Maturity Bid 10 Years, Thomson Reuters Source: Thomson Reuters Symbol: TRUS10C	CBOT
U.S. T-bond	US Treasury Bond (1987 8 ¾% - 2015) Source: NYSE Symbol: 747538	CBOT
British pounds	GBP to USD (BOE) Source: Bank of England Symbol: USSTBOE	CME
Euro	Euro to USD (BOE) Source: Bank of England Symbol: EUUSBOE	CME
Japanese yen	JPY to USD (BOE) Source: Bank of England Symbol: JPUSBOE	CME

3.2 Data on Shanker's (2017) indices of adequate and excess speculation and Working's (1960) speculative index T:

In the COT report, the CFTC classified traders' futures positions as commercial and non-commercial. In the literature, commercial positions have been treated as hedging and non-commercial positions have been treated as speculative. Shanker (2017) and Sanders et al. (2009), allocate non-reporting traders' open interest into commercial and non-commercial categories by assuming that the ratio of commercial to non-commercial position for non-reporting traders is the same as that for the reporting traders. Shanker (2017) estimated the indices of adequate and excess speculation and Working's speculative index T for the same 21 futures contracts for the period 1986-2015. Due to the release times of COT, INDADSP, INDEXSP, and T index are, for different futures contracts, are bi-weekly prior to September 1992, and weekly thereafter. I use these data (with permission from Professor Shanker) as measures of speculative activity in the 21 different futures markets in this thesis.

4. Methodology

4.1 Estimation of Expected and Unexpected Spot Market Volatility

I choose the most appropriate GARCH model from the standard GARCH (p, q), GJR-GARCH and EGARCH model to capture the underlying spot market volatilities for different commodities/financial assets. Following Chang et al. (2000), I decompose volatility into two components, expected volatility and unexpected volatility. A time series of daily spot market returns R_t is calculated for each of the 21 different commodities/financial assets as:

$$R_t = ln\left(\frac{p_t}{p_{t-1}}\right), \tag{4.1}$$

Where p_t represents the spot price of the commodity/financial asset on day t. In order to decide which model is better to use to estimate the spot price volatility for different underlying returns, I run the standard GARCH (combine the mean equation Eq**(4.4) (4.2)** and variance equation Eq**(4.3)**), GJR-GARCH (combine the mean equation Eq**(4.2)** and variance equation Eq**(4.4)**), and EGARCH model (combine the mean equation Eq**(4.2)** and variance equation Eq**(4.5)**) for all 21 commodities and financial assets. Following Pagan and Schwert (1990), Engle and Ng (1993) and Gulen and Mayhew (2000), I remove predictability associated with lagged returns from the time series of daily spot market returns. For each underlying asset, the mean equation is estimated in Eq**(4.2)**:

$$R_t = a + \sum_{i=1}^{n} b_i R_{t-i} + \varepsilon_t \tag{4.2}$$

The variance equations for standard GARCH, GJR-GARCH, and EGARCH model are estimated in Eq**(2.10)**, Eq**(4.4)** and Eq**(2.13)** respectively: :

$$EV_t^2 = \omega + \sum_{i=1}^{p} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j \, EV_{t-j}^2. \tag{4.3}$$

$$EV_t^2 = \omega + \sum_{i=1}^{p} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j \, EV_t^2{}_{t-j} + \theta S_{t-1}^- \varepsilon_{t-1}^2, \tag{4.4}$$

$$\ln(EV_t^2) = \omega + \frac{\alpha_1 \varepsilon_{t-1} + \gamma_1 |\varepsilon_{t-1}|}{EV_{t-1}} + \beta_1 \ln(EV_{t-1}^2). \tag{4.5}$$

where ω is the constant term, the conditional variance EV_t^2 depends on the GARCH specification and is a time-varying, positive and measurable function of past information, $S_{t-1}^- = 1$ if $\varepsilon_{t-1} < 0, S_{t-1}^- = 0$ otherwise. It can capture the impact of news on volatility and provide a good fit to the data. Since the probability distribution of asset returns often exhibit thicker tails and excess kurtosis than the standard normal distribution due to a volatility clustering in financial markets, in order to capture this phenomenon, the Student's t distribution is considered in my thesis. Therefore, ε_t is an error term and it assumed to follow either the normal or t distribution.

To compare the different models and test for whether these volatility models have adequately captured all of the persistence in the variance of returns, I apply several standard criteria: 1) the Ljung-Box Q^2 tests, $Q^2(20)$, which indicate whether there is autocorrelation at lag 20 of the squared standardized residuals. The null hypothesis is that there is no serial correlation in the residuals. If the model is adequate, then the standardized squared residuals should be serially uncorrelated. 2) ARCH-LM test which indicates whether there is an unexplained ARCH effect in the standardized residuals. The null hypothesis is that there is no ARCH effect. If the model is adequate, the null hypothesis should not be rejected. 3) Finally, following Sin and White (1996), I employ the AIC and Log-likelihood value to select the best model to describe the conditional dependence in the volatility process. The target is to choose the model with the lowest AIC value and the highest Log-likelihood value.

Following Chang et al. (2000), I further decompose volatility into two components, expected and unexpected volatility. Since $EV_t^2$is constructed based on the past information available at $t-1$, which can be known at the beginning of the trading day t, EV_t^2 is referred to as the expected conditional variance on day t. In order to make the following ARDL model more stable and eliminate the autocorrelation and heteroscedasticity, following Arize and Malindretos (1998), I scale the volatilities up by taking the logarithm of the square root of the conditional variance as the expected volatility (EV_t) at time t. Accordingly, the unexpected volatility (UEV_t) at time t is calculated as the logarithm of the square root of $|\varepsilon_t^2 - EV_t^2|$. Since the data on Shanker (2017)'s indices of adequate and excess speculation and Working's (1960) speculative index T are bimonthly prior to 1992 and weekly thereafter, for each date associated with the indices, following Shanker (2017), I find the corresponding value of weekly expected and unexpected volatility by averaging EV and UEV respectively, for a period following the pervious date and including the current date[2].

4.2 Estimation of the long run and short run relationship between futures market speculation and spot market volatility

I examine the stationarity of EV, UEV, INDADSP, INDEXSP and T using the Augmented Dickey-Fuller test (Dickey and Fuller, 1979) and use ARDL procedure to find the long run contemporary relationship between the variables. I use the AIC to select the maximum lag-length for all the variables in the ARDL model. Without knowing the direction of the long-run relationship between the variables, the following unrestricted error correction (UEC) regressions are estimated:

$$\begin{aligned} \Delta EV_t = \alpha_0 &+ \sum_{i=1}^{p} \beta_1 \Delta EV_{t-i} + \sum_{i=1}^{q} \beta_2 \Delta INDADSP_{t-i} + \sum_{i=1}^{m} \beta_3 \Delta INDEXSP_{t-i} \\ &+ \sum_{i=1}^{n} \beta_4 \Delta T_{t-j} + \tau_1 EV_{t-1} + \tau_2 INDADSP_{t-1} + \tau_3 INDEXSP_{t-1} \\ &+ \tau_4 T_{t-1} + e_t \end{aligned} \tag{4.6}$$

---

[2] e.g. $\sum_k \frac{EV_t}{K}$ and $\sum_k \frac{UEV_t}{K}$, where K is the number of days in the interval

$$
\begin{aligned}
\Delta UEV_t = \alpha_1 &+ \sum_{i=1}^{p} \beta_5 \Delta UEV_{t-i} + \sum_{i=1}^{q} \beta_6 \Delta INDADSP_{t-i} + \sum_{i=1}^{m} \beta_7 \Delta INDEXSP_{t-i} \\
&+ \sum_{i=1}^{n} \beta_8 \Delta T_{t-j} + \tau_5 UEV_{t-1} + \tau_6 INDADSP_{t-1} + \tau_7 INDEXSP_{t-1} \\
&+ \tau_8 T_{t-1} + e_t
\end{aligned}
\tag{4.7}
$$

where β_i represent the short-run coefficients and τ_i represents the long-run coefficients. α_i is the drift, e_t is the error term and Δ is the first difference operator. The maximum of lags, p, q, m, and n in **Eq(4.6)** and **Eq(4.7)** selected by AIC are used to determine the optimal structure for the ARDL (p, q, m, n) specification. The model with the smallest AIC and highest R^2performs relatively better.

The ARDL test for cointegration begins with the estimation of **Eq(4.6)** and **Eq(4.7)** using the ordinary least squares (OLS) method after selecting the optimal lag length. After the regressions of **Eq(4.6)** and **Eq(4.7)** were conducted, the Wald test (F-statistic) was used to test for the existence of a long-run relationship among the variables. The null hypothesis of non-existence of a long-run relationship and the alternative hypothesis are as follows:

$$H_0: \tau_1 = \tau_2 = \tau_3 = \tau_4 = 0.$$

$$H_1: \tau_1 \neq \tau_2 \neq \tau_3 \neq \tau_4 \neq 0.$$

The F-statistic was compared with Pesaran's (2001) critical values at different levels of significance. Since there are no exact critical values for F-test for an mix of I(0) and I(1) variables, the critical values proposed by Pesaran et al. (2001) have lower bound and upper bound for the asymptotic distribution of the F-statistic. The lower bound assumes that all the variables are I(0), meaning that there is no cointegration among the underlying variables and the upper critical bound assumes that all the variables are I(1), meaning that there is cointegration among the underlying variables. When the computed F-statistic is greater than the critical value of the upper bound, then the null is rejected.

If a stable long-run relationship is supported by the last step, in this step, the long-run coefficients of the dependent variable and the associated ARDL error correction models are estimated. First, the augmented ARDL (p, q, m, n) long-run model is estimated as follows:

$$EV_t = \alpha_0 + \sum_{i=1}^{p} \tau_1 EV_{t-i} + \sum_{i=1}^{q} \tau_2 INDADSP_{t-i} + \sum_{i=1}^{m} \tau_3 INDEXSP_{t-i} + \sum_{i=1}^{n} \tau_4 T_{t-j} + u_t \tag{4.8}$$

$$UEV_t = \alpha_0 + \sum_{i=1}^{p} \tau_5 UEV_{t-i} + \sum_{i=1}^{q} \tau_6 INDADSP_{t-i} + \sum_{i=1}^{m} \tau_7 INDEXSP_{t-i} + \sum_{i=1}^{n} \tau_8 T_{t-j} + u_t \tag{4.9}$$

The p, q, m, and n in **Eq(4.8)** and **Eq(4.9)** are the maximum of the lags selected by the AIC and they are the same as the number of lags in **Eq(4.6)** and **Eq(4.7).** The short run dynamic coefficients associated with the long-run relationships are estimated using an error correction model (ECM) presented in Eq**(4.10)** and Eq**(4.11)**.

$$\Delta EV_t = \alpha_1 + \sum_{i=1}^{p} \beta_1 \Delta EV_{t-i} + \sum_{i=1}^{q} \beta_2 \Delta INDADSP_{t-i} + \sum_{i=1}^{m} \beta_3 \Delta INDEXSP_{t-i} + \sum_{i=1}^{n} \beta_4 \Delta T_{t-j} + \omega ETC_{t-1} + \varepsilon_t \tag{4.10}$$

$$\Delta UEV_t = \alpha_1 + \sum_{i=1}^{p} \beta_5 \Delta UEV_{t-i} + \sum_{i=1}^{q} \beta_6 \Delta INDADSP_{t-i} + \sum_{i=1}^{m} \beta_7 \Delta INDEXSP_{t-i} + \sum_{i=1}^{n} \beta_8 \Delta T_{t-j} + \omega ETC_{t-1} + \varepsilon_t \tag{4.11}$$

The ETC_{t-1} term can be estimated using OLS by rearranging the original equation Eq **(4.8)** and Eq **(4.9)**. Under the ARDL approach, the existence of a unique valid long run relationship among the variables, and hence a sole error-correction term, Eq **(4.12)** is the basis for estimation and inference. This is specified as follows:

$$ETC_{t-1} = EV_{t-1} - \sum_{i=1}^{q} \tau_2 INDADSP_{t-i} - \sum_{i=1}^{m} \tau_3 INDEXSP_{t-i} - \sum_{i=1}^{n} \tau_4 T_{t-j} - \alpha_0 \tag{4.12}$$

Moreover, Pesaran and Pesran (1997) used cumulative sum (CUSUM) and the cumulative sum of squares (CUSUMQ) test introduced by Brown et al. (1975) to test the above long-run model for the stability of its parameters.

4.3 Estimation of a causal relationship between futures market speculation and spot market volatility

Following the methodology of Toda and Yamamoto (1995), Lehecka (2013), and Anguibi (2015), firstly, I determine the order of integration of each of the time-series variables by using the ADF test and set the maximum order of integration for the time-series to be d_{max}. Then I set up a vector autoregressive (VAR) model by determining the optimal lag order, say k, using the AIC. Next, I take the VAR (k) model (without autocorrelation in the residuals) and add in d_{max} additional lags of each of the variables into each of the equations, to estimate VAR ($k + d_{max}$), represented by the following Eq**(4.13)** and Eq**(4.14)**:

$$\begin{aligned} EV_t = \alpha_3 &+ \sum_{i=1}^{k} \tau_i EV_{t-i} + \sum_{i=1}^{k+d_{max}} \tau_i EV_{t-i} + \sum_{i=1}^{k} \beta_i INDADSP_{t-i} \\ &+ \sum_{i=1}^{k+d_{max}} \beta_i INDADSP_{t-i} + \sum_{i=1}^{k} \gamma_i INDEXSP_{t-i} \\ &+ \sum_{i=1}^{k+d_{max}} \gamma_i INDEXSP_{t-i} + \sum_{i=1}^{k} \delta_i T_{t-j} + \sum_{i=1}^{k+d_{max}} \delta_i T_{t-j} + u_t \end{aligned} \tag{4.13}$$

$$\begin{aligned} UEV_t = \alpha_4 &+ \sum_{i=1}^{k} \epsilon_i UEV_{t-i} + \sum_{i=1}^{k+d_{max}} \epsilon_i UEV_{t-i} + \sum_{i=1}^{k} \theta_i INDADSP_{t-i} \\ &+ \sum_{i=1}^{k+d_{max}} \theta_i INDADSP_{t-i} + \sum_{i=1}^{k} \rho_i INDEXSP_{t-i} \\ &+ \sum_{i=1}^{k+d_{max}} \rho_i INDEXSP_{t-i} + \sum_{i=1}^{k} \omega_i T_{t-j} + \sum_{i=1}^{k+d_{max}} \omega_i T_{t-j} + u_t \end{aligned} \tag{4.14}$$

where all variables are as previously defined in section 4.2. I estimated **Eq(4.13)** considering $EV_t, UEV_t, INDADSP_t, INDEXSP_t$, and T_t in turn as the dependent variable. The procedure is the same for **Eq(4.14).** The null hypotheses are that the coefficients of the lagged values of volatility, INDADSP, INDEXSP and T are zero respectively, using a modified Wald (MWALD) test. In this case, the MWALD test statistic will be asymptotically chi-squared with k degrees of freedom under the null hypothesis that expected volatility is not caused by speculation. For example, if EV is the dependent variable as specified in **Eq(4.13),** the null and alternate hypotheses are as shown below:

(i) INDADSP does not Granger-cause EV:

$$\begin{gathered} H_0: \beta_i = 0, \forall\ i < k \\ H_1: \beta_i \neq 0, \forall\ i < k \end{gathered} \tag{4.15}$$

(ii) INDEXSP does not Granger-cause EV:

$$\begin{gathered} H_0: \gamma_i = 0, \forall\ i < k \\ H_1: \gamma_i \neq 0, \forall\ i < k \end{gathered} \tag{4.16}$$

(iii) T does not Granger-cause EV:

$$\begin{gathered} H_0: \delta_i = 0, \forall\ i < k \\ H_1: \delta_i \neq 0, \forall\ i < k \end{gathered} \tag{4.17}$$

5. Empirical Analysis and Results:

This section details the steps of the empirical analysis and results for gold as an example of the 21 commodities/financial assets in sections 5.1 and 5.2 in two stages. The first stage addresses the empirical analysis and the results of the estimation of expected and unexpected volatilities from daily spot prices using the GARCH family of models. The second stage addresses the empirical analysis and the results of the estimation of the relationship between futures market speculation, as captured by the index of adequate speculation, the index of excess speculation and Working's T. The analyses of the rest of the commodities/financial assets follow the same procedure and the results are presented in section 5.3.

5.1 Results of the empirical analysis for gold:

5.1.1 Characteristics of spot prices and returns

Fig. 5.1.1 graphs the daily gold price over the period of analysis 1986-2015. This figure clearly shows that gold prices increased in periods bearish stock markets (1987 and 2008) and decreased in bullish stock markets (1996, 2006 and after 2012). This distinct cyclical behavior is shown clearly in the graph of the daily gold return over time, which is presented in **Fig. 5.1.2**. In **Fig. 5.1.2**, high volatility periods are indicated by dark circles. The graph illustrates the clustering behavior of volatility and the presence of ARCH effects, under which periods of high volatility tend to follow periods of high volatility and periods of low volatility tend to follow periods of low volatility.

Fig. 5.1.1 Daily gold prices

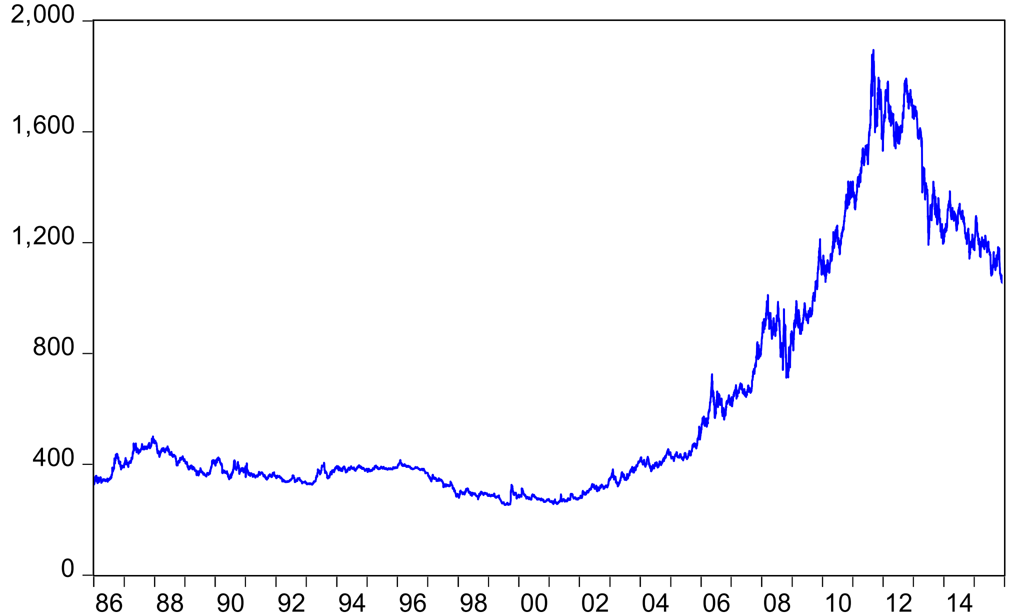


Fig. 5.1.2 Daily gold returns

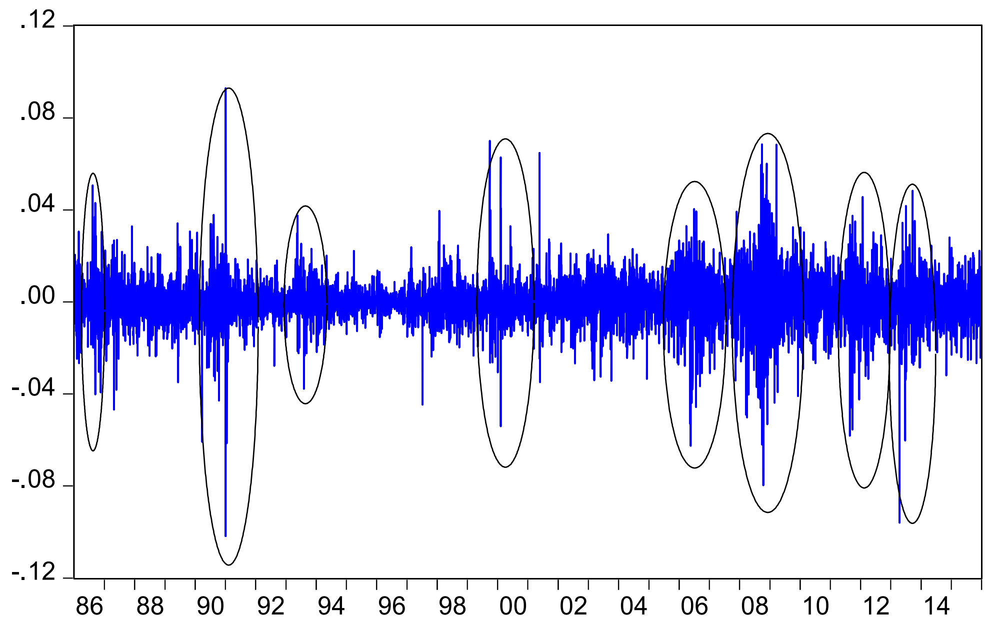


Table 5.1.1 presents summary statistics for daily gold spot returns (in Panel A), tests for heteroscedasticity and autocorrelation (in Panel B), as well as unit root test (in Panel C) in the same series, for the period January 3, 1986 through December 31, 2015. Panel A reveals that gold returns exhibit a very small mean.

This series also displays negative skewness and leptokurtic behavior, symptomatic of a heavier tailed distribution than the normal. The Jarque-Bera (J-B) test further confirms departure from normality of the daily returns. Therefore, the GARCH model with a Student's t distribution is considered in this case. Both the Ljung-Box (Q (20)) and the Breusch-Godfrey LM tests show serial dependence in the return series while the ARCH-LM and $Q^2(20)$ test in Panel B show evidence of heteroscedasticity. Panel C shows the results of the ADF unit root test of nonstationarity in the daily gold spot returns. The t-statistic indicates that at the 1% level of significance, the null hypothesis of nonstationarity is rejected. Therefore, the conclusion is that the daily returns for gold follow a stationary distribution.

Table 5.1.1 Descriptive statistics and unit root tests

Panel A. Summary Statistics						
Mean	Std. Dev.	Skewness	Kurtosis	J-B	Q(20)	LM(20)
0.000	0.010	-0.257	11.853	25645.030	52.488	2.664
Panel B. Heteroscedasticity and Autocorrelation tests						
ARCH-LM (p-value)				Q^2 (20) (p-value)		
416.86 (0.000)***				1791.6 (0.000)***		
Panel C. Unit roots tests (ADF)						
Test critical values				t-statistic		
1% level		-2.565		-91.999***		
5% level		-1.941				
10% level		-1.617				

*Notes: The table reports descriptive statistics and results of tests of heteroscedasticity and nonstationarity of daily gold spot returns over January 3, 1986 through December 31, 2015. J–B is the Jarque and Bera test for normality. I used 20 lags for both series. LM(20) refers to the Breusch–Godfrey test for the null hypothesis of no serial correlation in daily returns up to 20 lags. ARCH-LM denotes the ARCH test for the null of no autoregressive conditional heteroscedasticity up to 20 lags. The Ljung-Box statistic $Q^2(20)$ checks for serial correlation in the squared return series up to the 20th order. ADF denotes the Augmented Dickey and Fuller — ADF test for the null hypothesis of non-stationarity in daily returns. *** Indicates a rejection of the null hypothesis at the 1% significance level.*

5.1.2 Selection of the appropriate model from the GARCH family of models

Since there is serial correlation and heteroscedasticity in the daily returns, the GARCH family of models are considered. The basic estimation can be divided into two parts, one for the mean, which is a simple autoregressive AR (1) model and another for the variance which is identified by a particular GARCH specification, i.e., GARCH (1, 1), GJR-GARCH, and EGARCH with normal and Student's t distributions. **Table 5.1.2** compares the results of the estimation for the different models. The p-values associated with the $Q^2(20)$ and ARCH-LM for the GARCH (1,1) model and the GJR-GARCH model with either the normal distribution or the Student's t distribution are all larger than 10%, indicating that the null hypothesis of no autocorrelation and no serial correlation in the residuals cannot be rejected. However, the results for the EGARCH models with either the normal or t distribution all indicate that there still remains serial correlation in the residuals, which is not desirable. The GJR-GARCH (1, 1) model with a t distribution has the smallest (more negative) AIC value and the highest Log-likelihood value. Thus, to the conclusion is that the GJR-GARCH (1, 1) model with a t distribution specification (GJR-GARCH (1, 1)) is the most appropriate model to capture the volatility of daily gold spot returns.

Table 5.1.2 Results of the estimation of the GARCH Family of models for daily gold spot returns

Gold	Normal distribution			Student's t distribution		
	GJR(1,1)	GARCH(1,1)	EGARCH	**GJR(1,1)**	GARCH(1,1)	EGARCH
$Q^2(20)$	21.324 (0.378)	22.657 (0.306)	52.957 (0.000)	**20.151** **(0.449)**	22.191 (0.330)	54.613 (0.000)
ARCH-LM	1.044 (0.405)	1.101 (0.340)	2.614 (0.000)	**0.982** **(0.481)**	1.075 (0.369)	2.673 (0.000)
AIC	-6.624	-6.619	-6.622	**-6.770**	-6.765	-6.772
Log-Likelihood	25922.070	25899.450	25914.450	**26490.600**	26475.460	26502.140

The estimated coefficients of the GJR-GARCH model with a Student's t distribution are shown in **Table 5.1.3**. In general, all coefficients are found to be highly significant, except for the constant, α, which is not significant at any of the standard levels. A more in-depth analysis shows that the sign of the asymmetry parameter θ is negative indicating that there exist no leverage effects. In addition, the model predicts that good news has an impact on volatility more than bad news. The effect of good news on conditional variance is 2.119 times[3] more than bad news. Moreover, $\alpha + \beta$ of the GJR-GARCH (1, 1)-t model is close to 1, which suggests high persistence of volatility over time.

[3] Recall in section 2.2, the effect of good news on conditional variance is α and the effect of bad news on conditional variance is $(\alpha + \theta)$. Therefore, 0.089 / (0.089-0.047) = 2.119.

Table 5.1.3 GJR-GARCH (1, 1)-t Model for Gold Spot Returns

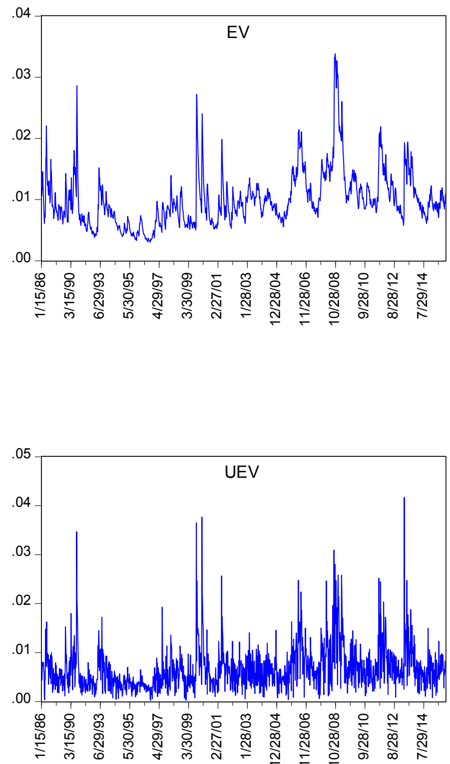
	Coefficient	Std. Error	t-Statistic	Prob.
Mean Equation				
a	0.000	0.000	1.171	0.242
b	-0.052	0.011	-4.884***	0.009
Variance Equation				
ω	0.000	0.000	3..470***	0.001
α	0.089	0.009	10.212***	0.000
β	0.942	0.004	217.553***	0.000
θ	-0.047	0.009	-5.281***	0.000
Student-t	3.690	0.185	19.905***	0.000

, **, * statistically significant at the 10%, 5% and 1% level of significance.[4]*

The expected volatility at time t (EV_t) is estimated as the logarithm of the square root of the conditional variance, as given by Eq**(4.4)**. The unexpected volatility at time t (UEV_t) is calculated as the logarithm of the square root of $|\varepsilon_t^2 - EV_t|$, where ε_t is defined by Eq**(4.2)** and Eq**(4.4)** together as an error term. The graphs of EV and UEV of daily gold spot returns are shown in **Fig 5.1.3**.

[4] The mean equation is $R_t = a + \sum_{i=1}^{n} b_i R_{t-i} + \varepsilon_t$ and the variance equation is $EV_t^2 = \omega + \sum_{i=1}^{p} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j \, EV_t^2{}_{t-j} + \theta S_{t-1}^{-} \varepsilon_{t-1}^2$ for GJR-GARCH model.

Fig 5.1.3 Expected and Unexpected Volatility of Daily Gold Spot Returns



5.2 Relationship between futures market speculation and spot market volatility for gold:

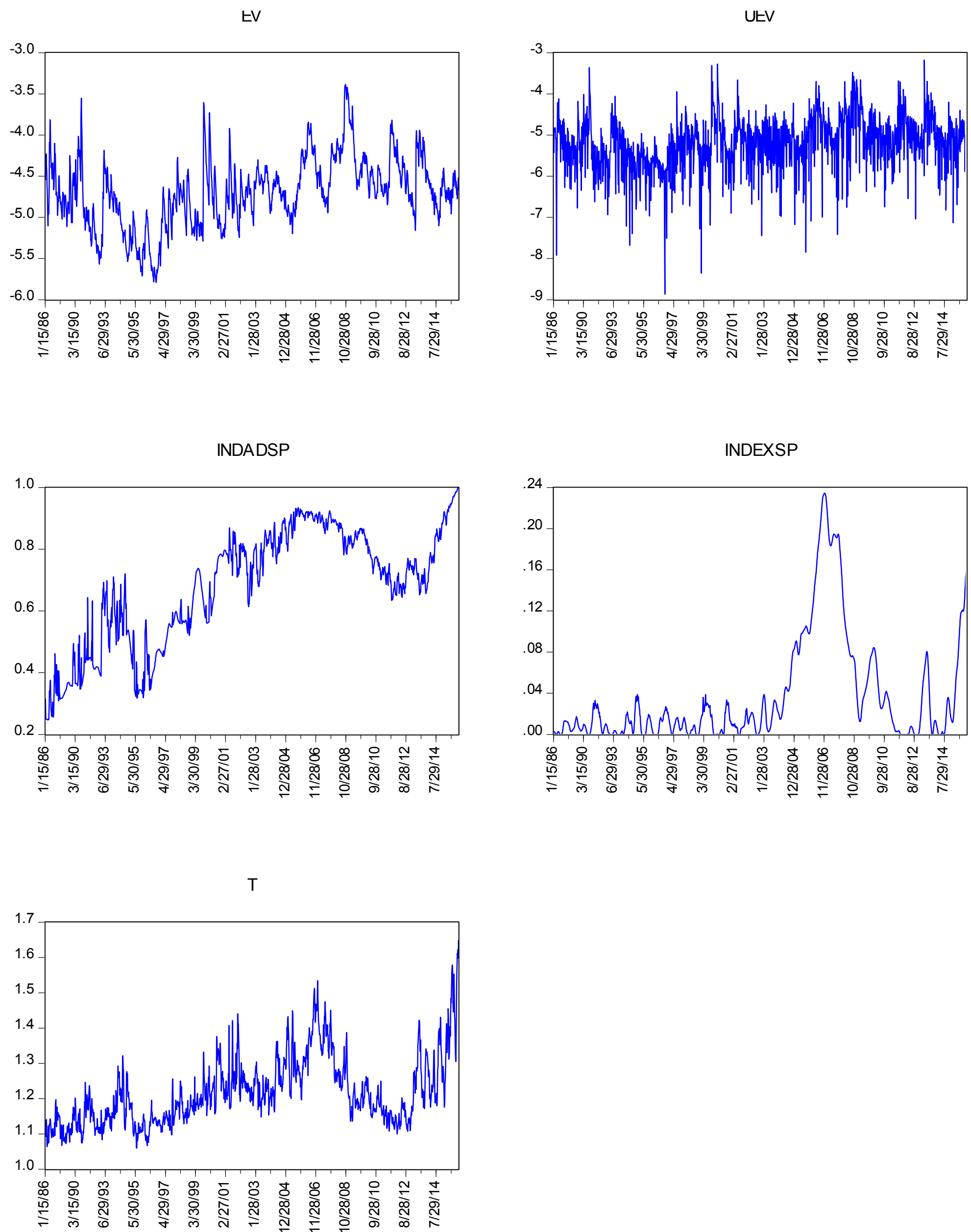
5.2.1 Characteristics of the variables used in the analysis

Next, I match the volatility and the speculative indices by week, as I stated in **Section 4.1**. The descriptive statistics of weekly values of these variables for the period January 15, 1986 through December 29, 2015 are shown in **Table 5.2.1.** The kurtosis values are all greater than three, except for INDADSP, indicating the presence of excess kurtosis in EV, UEV, INDEXSP and T. Excess skewness is also observed in UEV, INDADSP INDEXSP and T. UEV and INDADSP exhibit negative skewness, while INDEXSP and T exhibit positive skewness. The high valued Jarque-Bera statistics indicate non-normal distributions of all variables except EV.

Table 5.2.1 Descriptive statistics of EV, UEV, INDADSP, INDEXSP, and T

	EV	UEV	INDADSP	INDEXSP	T
Mean	-4.697	-5.200	0.678	0.037	1.218
Median	-4.700	-5.159	0.721	0.015	1.196
Maximum	-3.386	-3.178	1.000	0.234	1.649
Minimum	-5.791	-8.867	0.247	0.000	1.060
Std. Dev.	0.420	0.668	0.188	0.053	0.097
Skewness	0.084	-0.615	-0.495	1.989	1.109
Kurtosis	3.224	4.629	2.133	6.306	4.312
Jarque-Bera	4.477(0.107)	238.513(0.000)	99.308(0.000)	1532.642(0.000)	380.649(0.000)
Observations	1374	1374	1374	1374	1374

Fig 5.2.1 Weekly values of EV, UEV, INDADSP, INDEXSP, and T for the period January 15, 1986 through December 29, 2015



5.2.2 Analysis of Stationarity of Each Relevant Variable

I examine the stationarity of EV, UEV, INDADSP, INDEXSP and T using the ADF unit root test. Error! Reference source not found. shows the results.

Table 5.2.2 Results of the Augmented Dickey-Fuller Unit Root Test for Gold

Variable	Level		1st Difference		Level of integration
	No Trend	With Trend	No Trend	With Trend	
Expected Volatility (EV)	-5.676***	-6.403***	-26.354***	-26.345***	I(0)
Unexpected Volatility (UEV)	-8.258***	-9.011***	-21.761***	-21.753***	I(0)
INDADSP	-2.489	-3.807**	28.834***	-28.836***	I(0)
INDEXSP	-1.489	-1.880	-9.984***	-10.005***	I(1)
T	-4.614***	-6.174***	-18.298***	-18.307***	I(0)

, **, * statistically significant at the 10%, 5% and 1% level of significance. The null hypothesis is that the series is non-stationary*

Error! Reference source not found. shows that the ADF test statistics for EV and UEV are both significant at the 1% level, indicating that they are stationary series. INDADSP and T are non-stationary when no correction for trend is used, but trend stationary at the 1% and 5% level of significance respectively. However, Error! Reference source not found. shows INDEXSP is non-stationary at the level with and without correction for trend. After first differencing, ADF values for all variables are significant at the 1% level. Therefore, the results imply that INDEXSP is integrated of order one I(1), while the other variables are integrated of order zero I(0). The mixed results of the unit root tests, that the variables do not exhibit the same order of integration, imply the absence of a relationship between levels of these variables.

5.2.3 Cointegration test: Autoregressive-Distributed Lag (ARDL) test

Since the variables exhibit different orders of integration, the ARDL model is used, as it was suggested in section 2.3. Next, I use the AIC to select the maximum lag-length for all the variables in the ARDL model, presented in Eq**(4.6)** and Eq**(4.7)**. The results shown in **Fig 5.2.2** suggest that the best ARDL (p, q, m, n) model for **Eq(4.6)** is ARDL (4, 4, 7, 4) and the best ARDL model for **Eq**(**4.7)** is ARDL (8, 0, 1, 6). The best-performing model also provides the estimates of the associated long run coefficients, as indicated in Eq**(4.8)** and Eq**(4.9)**, and the Error Correction Model (ECM), as indicated in Eq**(4.10)** and Eq**(4.11)**.

Fig 5.2.2 Lag Specification by the Akaike Information Criteria (AIC)

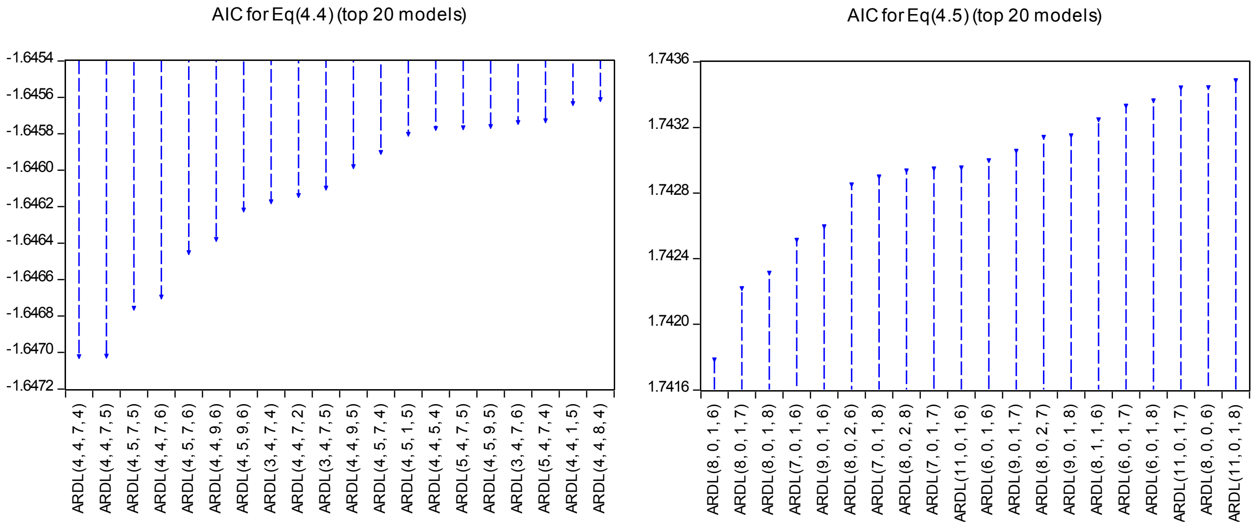


Table 5.2.3 shows the results of the estimation of **Eq(4.6)**. The results indicate that the lower bound value is 4.29, and the upper bound value is 5.61, at the 1% level of significance. The F-statistic (9.222) is significantly higher than both the upper and lower bounds, implying that the null hypothesis that no relationship between the variables is rejected. The results of the estimation of **Eq(4.7)** shown in the second row are similar. The conclusion is that there is cointegration between EV, INDADSP, INDEXSP, and T, as well as between UEV, INDADSP, INDEXSP and T. Once it is established that a long-run cointegrating relationship exists between the variables, the marginal impacts of INDADSP, INDEXSP and T on EV and UEV for gold are addressed.

Table 5.2.3 Results of the ARDL Bounds test for Eq (4.6) and Eq (4.7)

		1%		5%		10%		
Model	F-Statistic	Low	High	Low	High	Low	High	Cointegration
Eq(4.6) ARDL(4,4,7,4)	9.222***	4.29	5.61	3.23	4.35	2.72	3.77	Present
Eq(4.7) ARDL(8,0,1,6)	16.830***	4.29	5.61	3.23	4.35	2.72	3.77	Present

, **, * statistically significant at the 10%, 5% and 1% level of significance. The null hypothesis is that no long-run relationship exists between the two variables.*

5.2.4 Testing for Long and Short-Run Coefficients

Table 5.2.4, Panels A and B, present the results of the estimation of **Eq(4.8)** and **Eq(4.9)** with expected volatility and unexpected volatility as the dependent variable respectively. In the long run, there is a positive and significant relationship between INDADSP and the expected and unexpected volatility of gold spot returns. A 1% increase in INDADSP has related to about 1.537% increase in EV and 1.424% increase in

UEV, all other things being equal, which suggests that adequate speculation has positive relationship with expected volatility for gold in the long term. However, the long-term coefficients for INDEXSP and T, are not significant at the 1%, 5% and 10% levels.

Table 5.2.4 Estimated Long Run Coefficients using the ARDL Approach

Regressor	Coefficient	Standard Error	T-ratio (p-value)
Panel A: Equation 5 ARDL(4,4,7,4)			
INDADSP	1.537	0.510	3.012(0.002)***
INDEXSP	1.458	1.748	0.834(0.405)
T	-1.349	1.309	-1.030(0.303)
Panel B: Equation 6 ARDL(8,0,1,6)			
INDADSP	1.424	0.422	3.372(0.001)**
INDEXSP	0.807	1.487	0.543(0.587)
T	-1.361	1.091	-1.247(0.213)

, **, * statistically significant at the 10%, 5% and 1% level of significance. The ARDL is selected on the basis of the AIC.*

The short run dynamic coefficients associated with the long-run relationships are shown in **Table5.2.5**. The error-correction coefficient ω of ECT(-1) in **Eq(4.10)** and **Eq(4.11)** is negative (i.e. -0.048 in Panel A and -0.311 in Panel B), as required, and is statistically significant at the 1% level. This further confirms the existence of a stable long-run relationship between volatilities and speculation. Moreover, the coefficient of ECT(-1) represents the speed of adjustment of disequilibrium. Considering the results of the estimation of **Eq(4.10)**, the implication is that about 4.8% of any movements into disequilibrium caused by the previous year's shock are corrected for in the current year. In Panel B, the short-run coefficients of INDEXSP is statistically significant at the 5% level, although the long run coefficients presented in **Table 5.2.4** are not statistically significant.

Table5.2.5 Error Correction Representation for the Selected ARDL Model

Panel A: ARDL (4,4,7,4)				
Variable	**Coefficient**	**Std. Error**	**t-Statistic**	**Prob.**
C	-0.199	0.033	-6.081	0.000
D(EV(-1))	0.245	0.027	9.072	0.000
D(EV(-2))	-0.075	0.028	-2.717	0.007
D(EV(-3))	-0.046	0.027	-1.714	0.087
D(INDADSP)	1.026	0.126	8.152	0.000
D(INDADSP(-1))	0.151	0.129	1.170	0.242
D(INDADSP(-2))	-0.194	0.129	-1.506	0.132

D(INDADSP(-3))	0.375	0.130	2.895	0.004
D(INDEXSP)	-3.829	2.379	-1.610	0.108
D(INDEXSP(-1))	-1.813	2.626	-0.691	0.490
D(INDEXSP(-2))	-4.944	2.736	-1.807	0.071
D(INDEXSP(-3))	-0.602	2.746	-0.219	0.826
D(INDEXSP(-4))	5.913	2.500	2.365	0.018
D(INDEXSP(-5))	-4.502	2.441	-1.844	0.065
D(INDEXSP(-6))	4.729	2.196	2.154	0.032
D(T)	0.398	0.122	3.268	0.001
D(T(-1))	0.386	0.123	3.149	0.002
D(T(-2))	0.163	0.124	1.311	0.190
D(T(-3))	0.280	0.125	2.246	0.025
ETC(-1)	-0.048	0.008	-6.080	0.000

R-squared	0.155	Mean dependent var	0.000
Adjusted R-squared	0.143	S.D. dependent var	0.114
S.E. of regression	0.105	Akaike info criterion	-1.648
Sum squared resid	14.956	Schwarz criterion	-1.572
Log likelihood	1146.491	Hannan-Quinn criter.	-1.620
F-statistic	12.997	Durbin-Watson stat	2.004
Prob(F-statistic)	0.000		

Panel A: ARDL (8,0,1,6)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-1.412	0.173	-8.177	0.000
D(UEV(-1))	-0.482	0.041	-11.704	0.000
D(UEV(-2))	-0.368	0.042	-8.703	0.000
D(UEV(-3))	-0.245	0.042	-5.848	0.000
D(UEV(-4))	-0.190	0.040	-4.683	0.000
D(UEV(-5))	-0.149	0.038	-3.884	0.000
D(UEV(-6))	-0.078	0.034	-2.291	0.022
D(UEV(-7))	-0.047	0.027	-1.743	0.082
D(INDEXSP)	-17.474	8.473	-2.062	0.039
D(T)	-0.237	0.617	-0.383	0.702
D(T(-1))	-0.185	0.577	-0.321	0.748
D(T(-2))	0.669	0.574	1.166	0.244
D(T(-3))	0.443	0.567	0.782	0.435
D(T(-4))	-0.678	0.562	-1.206	0.228
D(T(-5))	1.907	0.565	3.375	0.001
ETC(-1)	-0.311	0.038	-8.214	0.000

R-squared	0.399	Mean dependent var	0.000
Adjusted R-squared	0.392	S.D. dependent var	0.742
S.E. of regression	0.578	Akaike info criterion	1.754
Sum squared resid	451.191	Schwarz criterion	1.815

Log likelihood	-1181.676	Hannan-Quinn criter.	1.776
F-statistic	59.726	Durbin-Watson stat	2.009
Prob(F-statistic)	0		

, **, * statistically significant at the 10%, 5% and 1% level of significance. ARDL selected on the basis of AIC*

5.2.5 Diagnostic

In order to address the unbiasedness and consistency of the estimated coefficients, following Ahmed (2013), I conduct three diagnostic tests: (A) the Breusch-Godfrey Serial Correlation LM test (B) the heteroscedasticity test based on the regression of squared residuals on squared fitted values. (C) the Ramsey Regression Equation Specification Error Test (RESET) test using the square of fitted values. The results are shown in **Table 5.2.6.**

Table 5.2.6 Diagnostic tests

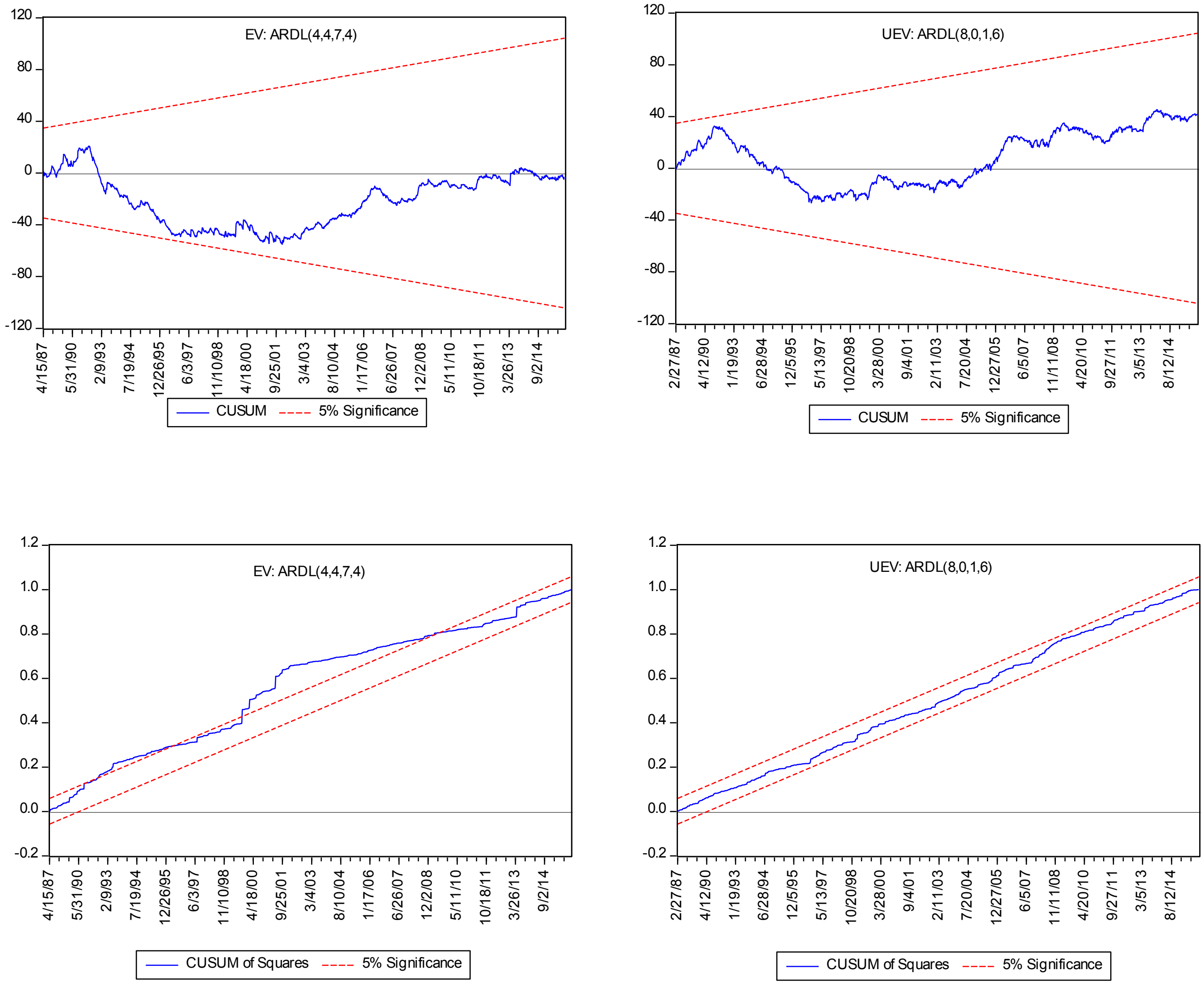
Equation	Serial Correlation (A) F-statistic (p-value)	Heteroscedasticity (B) F-statistic (p-value)	Functional Form (c) F-statistic (p-value)
Eq. ARDL(4,4,7,4)	F(15,1329)=1.242 (0.233)	F(22,1344)=2.902 (0.000)	F(2,1342)=1.386 (0.251)
Eq. ARDL(8,0,1,6)	F(15,1332)=1.094 (0.356)	F(18,1347)=1.293 (0.182)	F(2,1345)=0.547 (0.579)

*, **, *** statistically significant at the 10%, 5% and 1% level of significance. The nulls are (A) No serial correlation (B) No Heteroscedasticity in the error terms (C) No functional-form misspecification respectively

The results of the diagnostic tests indicate that the estimation of long-run coefficients and the ECM are free from serial correlation and functional-form misspecification at all 1%, 5% and 10% levels of significance for both equations. Although (B) test shows that we cannot reject the null hypothesis that there is no heteroscedasticity in the error terms, according to Shresthe and Chowdhury (2005), "since the time series constituting the ARDL equation are potentially of mixed order of integration, i.e., I(0) and I(1), it is natural to detect heteroscedasticity", I do not correct for heteroscedasticity.

As is suggested in section 4.3, plots of the cumulative sum of recursive residuals (CUSUM) and cumulative sum of squares (CUSUMSQ), based on the residuals from a sequential regression are shown in **Fig 5.2.3**. The plots of both CUSUM and CUSUMSQ indicate that the parameters of the long run equation over the sample period for the unexpected volatility model are stable. The red lines indicate the values of the critical bounds for CUSUM or CUSUMSQ at the 5% level of significance.

Fig 5.2.3 Plots of CUSUM and CUSUMSQ for the Coefficients of the ARDL Model



5.2.6 Toda-Yamamoto (T-Y) Causality Testing:

According to the stationarity specification in Error! Reference source not found., the order of integration of EV, UEV, INDADSP, INDEXSP, and T is I(0), I(0), I(0), I(1), and I(0) respectively. Therefore, the maximum order of integration for this group of time series is one, i.e. $d_{max} = 1$. Including the right number of lags in the VAR model should allow the model to be free from the effect of serial correlation. The dependent variable for VAR model A is EV, the dependent variable for Model B is UEV and the independent variables for both Models A and B are INDADSP, INDEXSP, and T. The results in **Table 5.2.7** suggest the use of 11 lags for both Model A and B. i.e. k=11. Therefore, VAR($k + d_{max}$) is 12.

Table 5.2.7 Lag – length selection for VAR model A and B

Lags	Model A			Model B		
	LM-Stat	Prob	AIC	LM-Stat	Prob	AIC
0	-	-	-8.155	-	-	-8.043
1	7.894	0.952	-22.298	8.019	0.948	-19.848
2	7.338	0.966	-23.269	7.707	0.957	-20.743
3	19.827	0.228	-23.430	11.728	0.763	-20.870
4	18.617	0.289	-23.443	14.307	0.576	-20.884
5	31.684	0.011	-23.442	24.258	0.084	-20.882
6	20.997	0.179	-23.437	14.499	0.562	-20.876
7	17.883	0.331	-23.433	19.521	0.243	-20.876
8	9.721	0.881	-23.468	9.997	0.867	-20.906
9	21.054	0.176	-23.461	19.947	0.223	-20.897
10	13.386	0.644	-23.487	14.291	0.577	-20.916
11	14.424	0.567	-23.495*	11.053	0.806	-20.920*

** indicates lag order selected by the criterion. The null hypothesis for the LM-test is that there is no serial correlation at lag order k.*

The empirical results of the Toda-Yamamoto causality test of **Eq(4.13)** and **Eq(4.14)** are summarized in **Table 5.2.8** and ***,** ***, *** statistically significant at the 10%, 5% and 1% level of significance. The null hypothesis test is that*

Table 5.2.9 respectively. The results reveal that there is bidirectional causality between the expected volatility of spot returns and INDADSP at the 10% level of significance. There is also bidirectional causality between INDEXSP and EV at the 10% level of significance. There is no causality relationship between EV and T. On the other hand, there is a one-way causal effect going from T to the unexpected volatility at the 5% level of significance. However, there is non-causality between unexpected volatility and INDADSP and INDEXSP. There is also no causal relationship between INDEXSP and INDADSP.

Table 5.2.8 Toda-Yamamoto Causality Test Results for Eq (4.13)

Dependent variable: EV			
H_0	Chi-sq	df	Prob.
INDADSP → EV	19.053	11	0.060*
INDEXSP → EV	23.429	11	0.015**
T → EV	13.608	11	0.255
INDADSP, INDEXSP, T → EV	46.162	33	0.064*
Dependent variable: INDADSP			
H_0	Chi-sq	df	Prob.
EV → INDADSP	23.390	11	0.016**
INDEXSP → INDADSP	6.032	11	0.871

T→ INDADSP	16.051	11	0.139
EV, INDEXSP, T → INDADSP	48.996	33	0.036**
Dependent variable: INDEXSP			
H_0	Chi-sq	df	Prob.
EV → INDEXSP	17.704	11	0.089*
INDADSP → INDEXSP	14.113	11	0.227
T → INDEXSP	191.920	11	0.000***
EV, INDADSP, T → INDEXSP	241.681	33	0.000***
Dependent variable: T			
H_0	Chi-sq	df	Prob.
EV → T	14.298	11	0.217
INDADSP → T	13.381	11	0.269
INDEXSP → T	136.703	11	0.000***
EV, INDADSP, INDEXSP → T	182.941	33	0.000***

, **, * statistically significant at the 10%, 5% and 1% level of significance. The null hypothesis test is that*

Table 5.2.9 Toda-Yamamoto Causality test Results for Eq (4.14)

Dependent variable: UEV			
H_0	Chi-sq	df	Prob.
INDADSP → UEV	14.007	11	0.233
INDEXSP → UEV	17.137	11	0.104
T → UEV	20.202	11	0.043**
INDADSP, INDEXSP, T → UEV	45.637	33	0.071*
Dependent variable: INDADSP			
H_0	Chi-sq	df	Prob.
UEV → INDADSP	16.914	11	0.110
INDEXSP → INDADSP	4.743	11	0.943
T → INDADSP	16.055	11	0.139
UEV, INDEXSP, T → INDADSP	41.593	33	0.145
Dependent variable: INDEXSP			
H_0	Chi-sq	df	Prob.
UEV → INDEXSP	11.998	11	0.364
INDADSP → INDEXSP	12.369	11	0.337
T → INDEXSP	201.388	11	0.000***
UEV, INDADSP, T → INDEXSP	235.451	33	0.000***
Dependent variable: T			
H_0	Chi-sq	df	Prob.
UEV → T	14.477	11	0.208
INDADSP → T	12.539	11	0.325
INDEXSP → T	146.690	11	0.000***

UEV, INDADSP, INDEXSP → T	183.144	33	0.000***

, **, * statistically significant at the 10%, 5% and 1% level of significance. The null hypothesis test is that*

5.3 Results of the empirical analysis for the 21 commodities/financial assets

Table 5.3.1 Descriptive Statistics and Unit Root Tests of daily spot returns for 21 commodities/financial assets

Commodity/ Financial Asset		Mean	Std. Dev.	Skewness	Kurtosis	J-B	Q(20)	LM(20)	Obs	ARCH-LM	Q^2(20)	ADF
Energy	Crude oil	0.000	0.026	-0.922	24.632	153691.400	106.260	5.738	7826	338.723***	2165.700***	-55.113***
	Natural Gas	0.000	0.038	0.540	14.986	40526.450	305.280	12.883	6716	345.580***	1971.500***	-24.049***
	Heating oil	0.000	0.025	-0.710	20.974	106004.400	31.414	1.593	7826	370.821***	1594.000***	-86.741***
Agricultural	Corn	0.000	0.018	-0.265	7.266	6025.542	37.687	1.874	7826	293.154***	2686.500***	-87.931***
	Soybean	0.000	0.015	-0.767	10.070	17067.400	40.084	1.954	7826	144.470***	1518.800***	-91.055***
	Wheat	0.000	0.022	-0.402	13.972	39465.320	57.977	2.889	7826	337.254***	906.910***	-94.086***
Metal	Copper	0.000	0.016	-0.229	7.870	7801.365	27.440	1.414	7826	396.952***	5102.700***	-90.766***
	Gold	0.000	0.010	-0.257	11.853	25645.030	52.488	2.664	7826	416.860***	1791.600***	-92.015***
	Sliver	0.000	0.018	-0.747	13.077	33841.540	26.718	1.350	7826	226.335***	1218.900***	-88.181***
Livestock	Lean hogs	0.000	0.015	-0.016	4.402	641.441	250.980	11.415	7826	551.520***	2607.300***	-54.510***
	Live cattle	0.000	0.009	-0.129	4.656	916.177	180.350	8.019	7826	466.017***	1649.300***	-36.817***
	Feeder	0.000	0.014	-0.008	8.589	7807.737	6611.4	114.233	5999	1418.365***	6645.400***	-12.186***
Currency	British pounds	0.000	0.006	0.145	6.524	4077.775	57.292	2.764	7826	130.186***	2555.100***	-83.689***
	Euro	0.000	0.006	-0.049	5.475	2000.015	24.050	1.187	7826	31.780***	967.830***	-86.765***
	Japanese yen	0.000	0.007	-0.411	7.677	7354.773	34.029	0.347	7826	180.113***	877.590***	-87.680***
Fixed-income	Eurodollar	0.000	0.107	0.052	23.786	140896.700	1579.500	154.969	7826	1425.462***	7462.600***	-36.885***
	10-Year T-note	0.000	0.015	-0.027	10.385	17784.970	59.013	3.006	7826	176.817***	3136.600***	-65.047***
	U.S. T-bond	0.000	0.031	1.230	27.225	193335.100	262.220	13.423	7826	160.378***	2374.400***	-17.871***
Stocks index	DJIA	0.000	0.011	-1.713	45.922	604573.400	60.519	3.044	7826	85.994***	830.950***	-67.070***
	NASDAQ	0.000	0.017	-0.094	10.794	19819.720	60.508	3.039	7826	547.103***	5699.800***	-65.719***
	S&P500	0.000	0.011	-1.289	31.535	267680.000	66.459	3.439	7826	148.980***	1601.800***	-66.837***

, **, * statistically significant at the 10%, 5% and 1% level of significance. The null hypothesis is that the series is nonstationary.*

The results indicate that the daily spot returns for all 21 commodities/financial assets have a zero mean. Most of the series of daily spot returns display negative skewness, except for Natural Gas, British Pounds, Eurodollar, and the U.S. T-bond. All series exhibit excess kurtosis (kurtosis is larger than 3). The Q(20) and LM(20) statistics indicate serial dependence and $Q^2(20)$ and ARCH-LM statistics indicate heteroscedasticity for all series. Hence, the GARCH family of models is used to model expected and unexpected volatility of the daily spot returns for all series. The ADF test statistics indicates that for all the series, the null hypothesis of non-stationarity is rejected. Since the results in the table above show non-standard distribution of the returns, the Student's t distribution is considered in the following for the GARCH family of models.

Table 5.3.2 Estimated Results of the Selected GJR model

Commodity/ Financial Asset		Coefficients of the Mean Equation			Coefficients of the Variance Equation					Diagnostic tests		
		C	AR(1)	AR(2)	ω	ε_{t-1}^2	ε_{t-2}^2	EV_{t-1}	$S_{t-1}^{-}\varepsilon_{t-1}^2$	Model	ARCH-LM	Q^2(20)
Energy	Crude oil	0.000	-0.030**		0.000***	0.077***	-0.020**	0.932***	0.017***		0.181	19.596
		(0.761)	(0.011)	-	(0.000)	(0.000)	(0.021)	(0.000)	(0.001)	GJR(2,1)	(0.671)	(0.419)
	Natural Gas	0.000	0.011***		0.000	0.106***	-0.087***	0.811***	5.381***		0.013	0.262
		(0.928)	(0.000)	-	(0.540)	(0.000)	(0.000)	(0.000)	(0.001)	GJR(2,1)	(1.000)	(1.000)
	Heating oil	0.000	-0.02		0.000***	0.131***	-0.037***	0.909***	-0.023***		1.334	27.118
		(0.075)	(0.863)	-	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	GJR(2,1)	(0.145)	(0.132)
Agricultural	Corn	0.000**	-0.004		0.000***	0.079***		0.916***	-0.001	GJR(1,1)	1.023	20.775
		(0.020)	(0.715)	-	(0.000)	(0.000)	-	(0.000)	(0.876)	-T dis	(0.430)	(0.411)
	Soybean	0.000***	-0.055***		0.000***	0.084***		0.936***	-0.046***	GJR(1,1)	0.843	12.826
		(0.000)	(0.000)	-	(0.000)	(0.000)	-	(0.000)	(0.000)	-T dis	(0.630)	(0.616)
	Wheat	0.000	-0.016		0.000***	0.069***		0.918***	-0.004	GJR(1,1)	1.297	25.761
		(0.881)	(0.142)	-	(0.000)	(0.000)	-	(0.000)	(0.700)	-T dis	(0.168)	(0.174)
Metal	Copper	0.000	-0.036***		0.000***	0.040***		0.955***	0.010	GJR(1,1)	1.001	20.187
		(0.359)	(0.001)	-	0.001)	(0.000)	-	(0.000)	(0.126)	-T dis	(0.457)	(0.446)
	Gold	0.000	-0.052***		0.000***	0.090***		0.942***	-0.048***	GJR(1,1)	0.990	20.302
		(0.242)	(0.000)	-	(0.000)	(0.000)	-	(0.000)	(0.000)	-T dis	(0.471)	(0.439)
	Sliver	0.000	-0.051***		0.000***	0.129***	-0.067***	0.964***	-0.045***	GJR(2,1)	0.781	15.747
		(0.090)*	(0.000)	-	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	-T dis	(0.740)	(0.732)
Livestock	Lean hogs	0.000	0.094***	0.050***	0.000***	0.082***	-0.060***	0.952***	0.032***	GJR(2,1)	1.324	27.857
		(0.804)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	-T dis	(0.151)	(0.113)
	Live cattle	0.000	0.061***		0.000***	0.019***		0.956***	0.037***	GJR(1,1)	1.461	28.345
		(0.302)	(0.000)	-	(0.000)	(0.000)	-	(0.000)	(0.000)	-T dis	(0.084)	(0.101)
Foreign Currency	British pounds	0.000	0.027**		0.000***	0.048***		0.955***	-0.015**	GJR(1,1)	1.326	27.954
		(0.130)	(0.017)	-	(0.000)	(0.000)	-	(0.000)	(0.020)	-T dis	(0.150)	(0.111)
	Euro	0.000	0.006		0.000***	0.040***		0.960***	-0.004	GJR(1,1)	1.257	25.597
		(0.716)	(0.604)	-	(0.000)	(0.000)	-	(0.000)	(0.460)	-T dis	(0.197)	(0.180)
	Japanese yen	0.000	-0.020*		0.000***	0.039***		0.941***	0.016**	GJR(1,1)	0.994	19.964
		(0.189)	(0.065)	-	(0.000)	(0.000)	-	(0.000)	(0.026)	-T dis	(0.466)	(0.460)
Fixed-income	Eurodollar	0.000***	-0.302***		0.000***	0.494***		0.769***	-0.050	GJR(1,1)	1.147	18.376
		(0.000)	(0.000)	-	(0.000)	(0.000)	-	(0.000)	(0.271)	-T dis	(0.295)	(0.497)
	10-Year T-note	0.000***	0.027**	-	0.000***	0.028***	-	0.961***	0.026***	GJR(1,1)	1.182	22.547

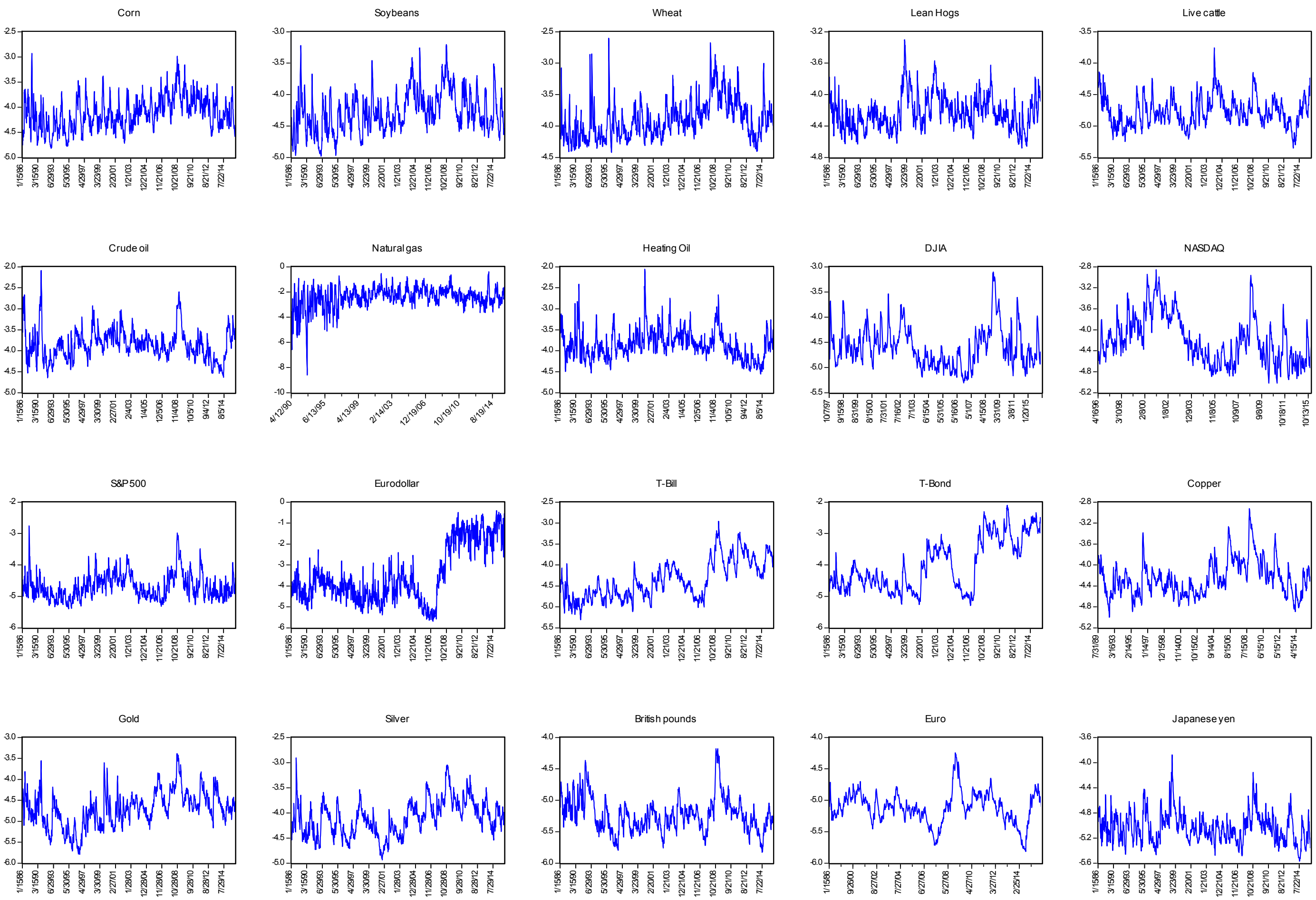
		(0.004)	(0.016)		(0.0639)	(0.000)		(0.000)	(0.000)	-T dis	(0.263)	(0.258)
	U.S. T-bond	0.000	0.015		0.000***	0.034***		0.959***	0.024***	GJR(1,1)	1.381	25.501
		(0.765)	(0.145)	-	(0.000)	(0.000)	-	(0.000)	(0.000)	-T dis	(0.124)	(0.145)
Stock index	DJIA	0.001***	-0.016		0.000***	0.016***		0.913***	0.110***	GJR(1,1)	0.470	9.413
		(0.000)	(0.176)	-	(0.000)	(0.017)	-	(0.000)	(0.000)	-T dis	(0.978)	(0.978)
	NASDAQ	0.001***	0.022**		0.000***	0.032***		0.916***	0.087***	GJR(1,1)	0.682	13.858
		(0.000)	(0.069)	-	(0.000)	(0.000)	-	(0.000)	(0.000)	-T dis	(0.848)	(0.838)
	S&P500	0.000***	-0.015		0.000***	0.007***		0.915***	0.129***	GJR(1,1)	0.409	8.233
		(0.000)	(0.221)	-	(0.000)	(0.305)	-	(0.000)	(0.000)	-T dis	(0.991)	(0.990)

, **, * statistically significant at the 10%, 5% and 1% level of significance. The value in parentheses represent p-values.*

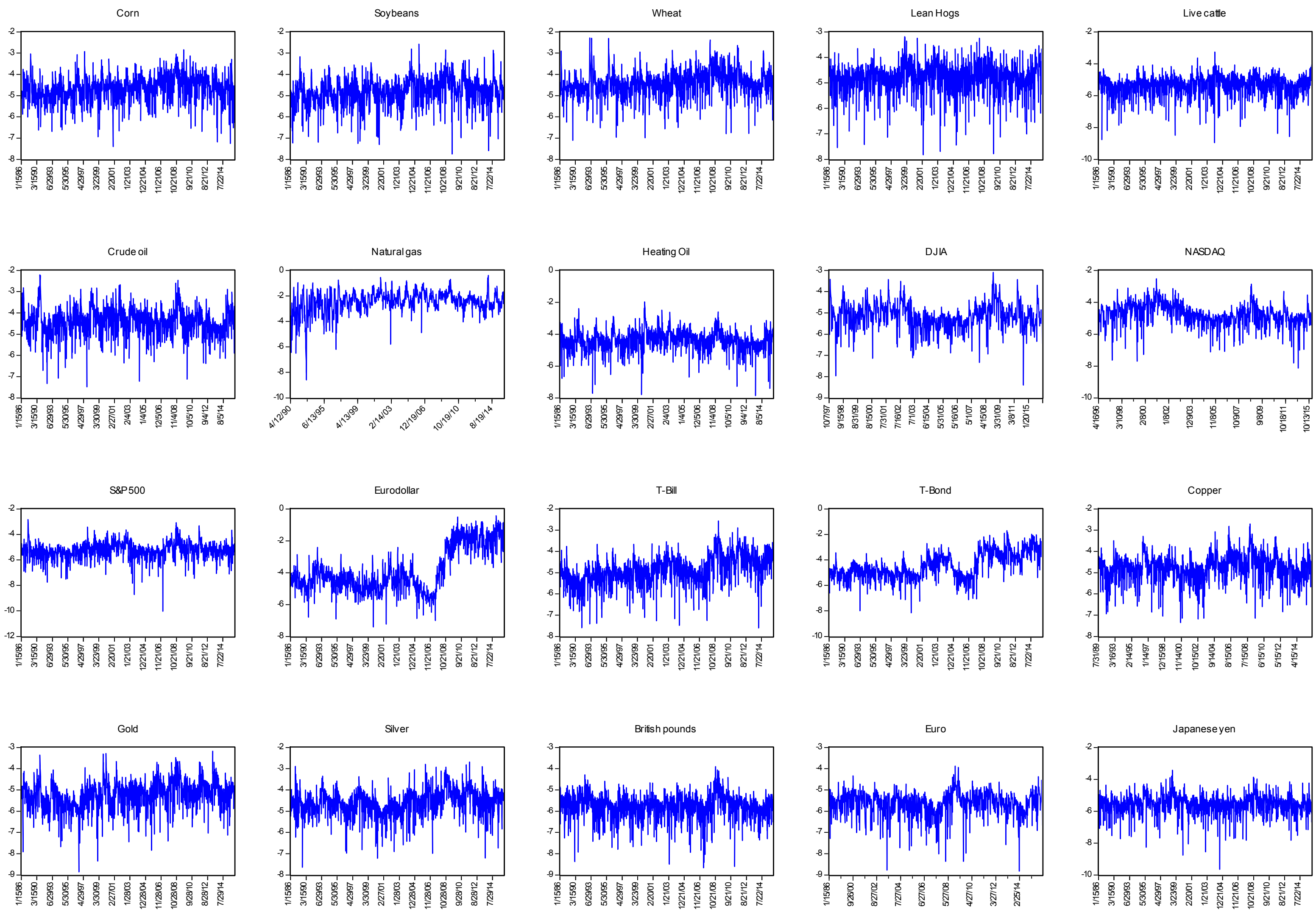
Table 5.3.2 reports the results of the estimation of the standard GARCH, GJR-GARCH and EGARCH models, along with the normal and Student's t distributions. The selected model is indicated in the column with heading "Model" showing that the volatilities of most of the series of daily spot returns can be best captured by the GJR-GARCH(1,1) model using the t distribution. However, for the energy commodities, volatilities are best modeled by the GJR-GARCH(2,1) model. The p-value associated with the ARCH-LM and $Q^2(20)$ statistics indicate lack of autocorrelation and serial correlation in the residual returns of 20 selected models, which is consistent with the conclusion that the selected model is correctly specified. However, the statistics of ARCH-LM and $Q^2(20)$ of feeder cattle indicate autocorrelation and serial correlation for all standard GARCH, GJR-GARCH, and EGARCH models with either normal or t-distribution. Since the conditional variance of feeder cattle under these model is not correctly specified, I drop it and only keep 20 commodities/financial assets in the following analysis. The coefficients of $S_{t-1}^{-}\varepsilon_{t-1}^{2}$ is significantly positive for crude oil, natural gas, all the livestock, Japanese yen, T-bond, T-note and all the stock index meaning that bad news have more impact on volatility than good news do. On the other hand, for heating oil, soybean, gold, silver and British pounds, good news bring more volatilities on spot price than bad news do. However, the effects of good news and bad news are not relevant on spot price volatility for corn, wheat, copper, euro, and Eurodollar.

Fig 5.3.1 Weekly data for 20 Commodities/Financial Assets

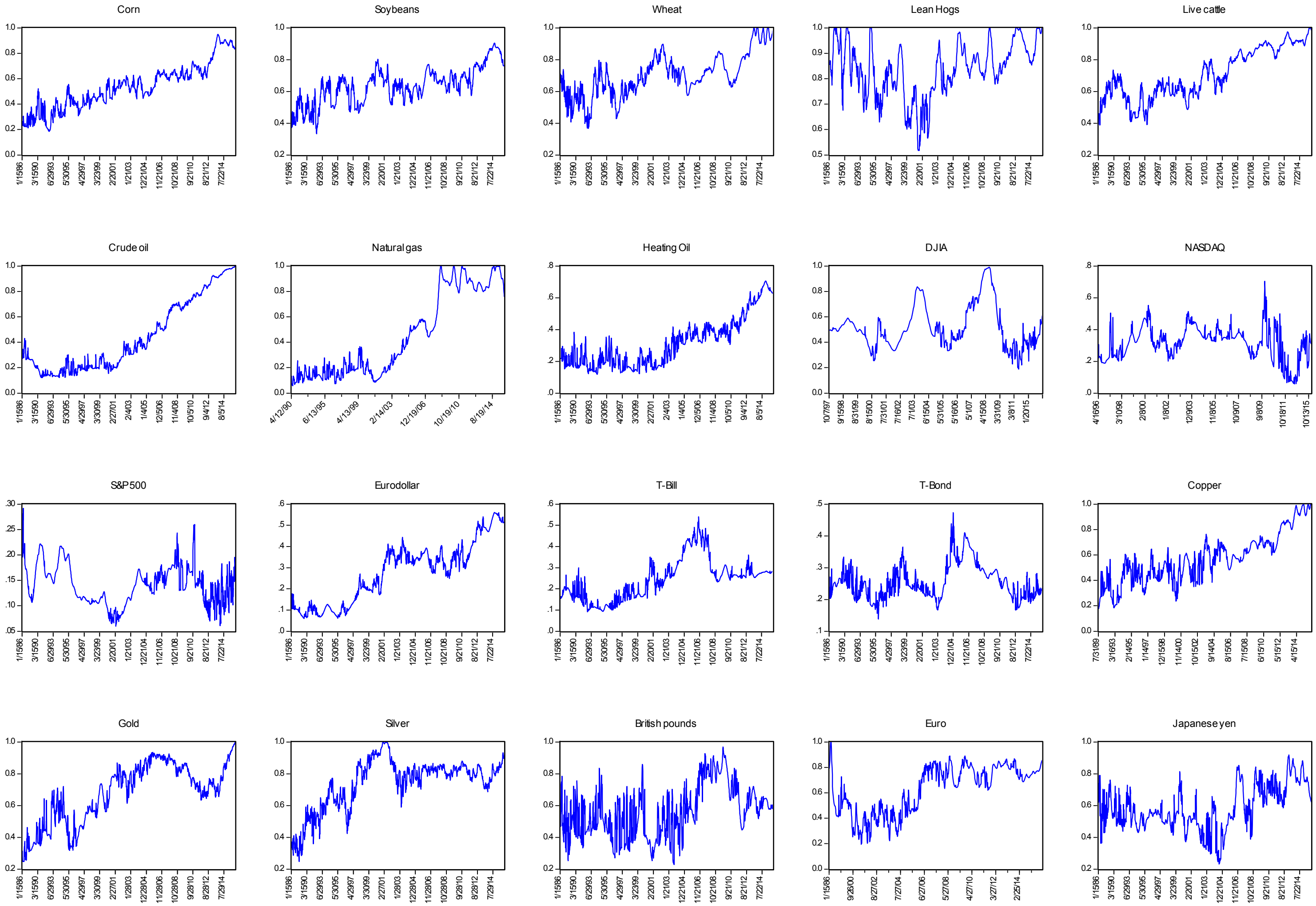
Weekly Expected Volatility for 20 Commodities/Financial Assets



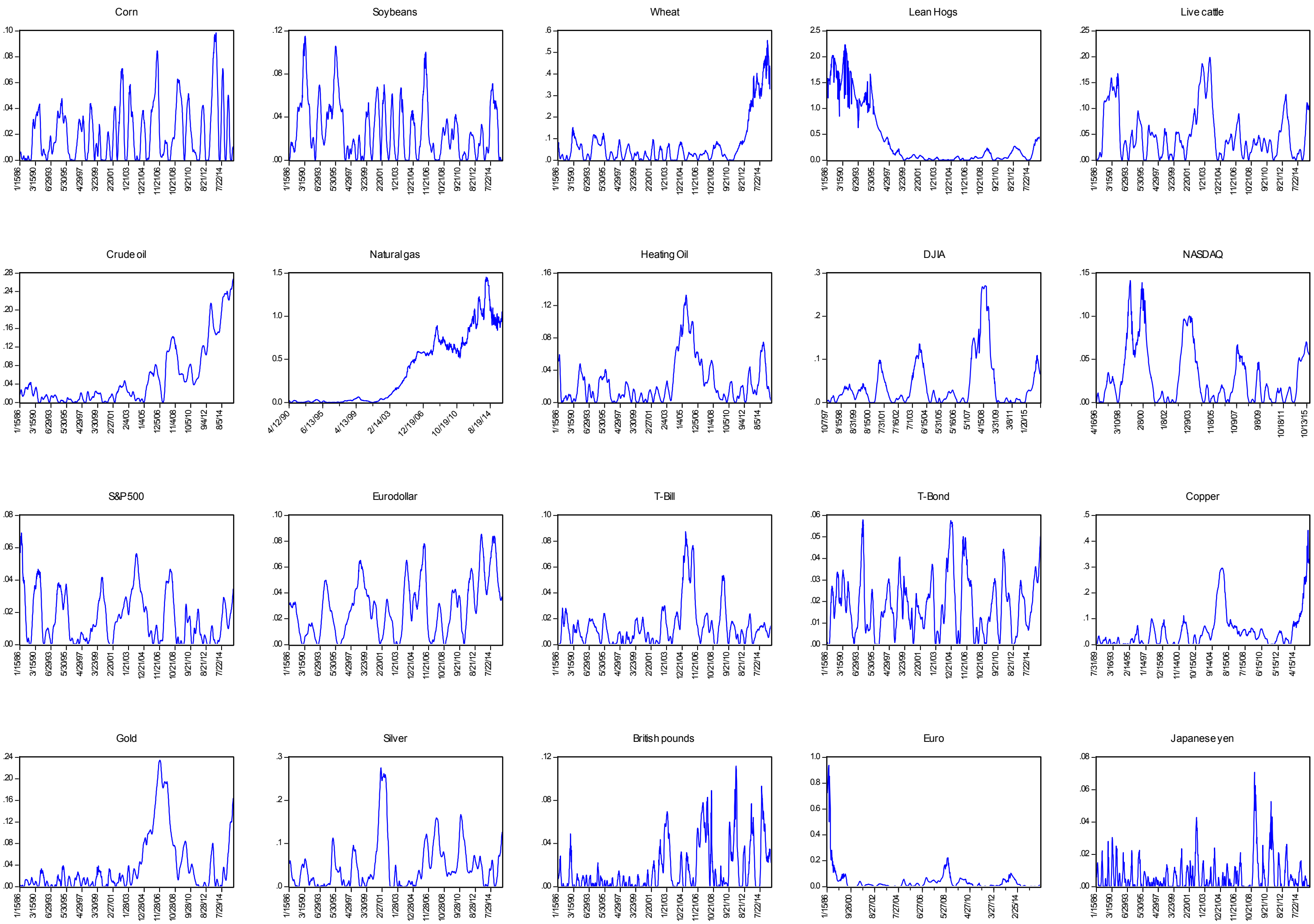
Weekly Unexpected Volatility for 20 Commodities/Financial Assets



Weekly INDADSP for 20 Commodities/Financial Assets



Weekly INDEXSP for 20 Commodities/Financial Assets



Weekly T for 20 Commodities/Financial Assets

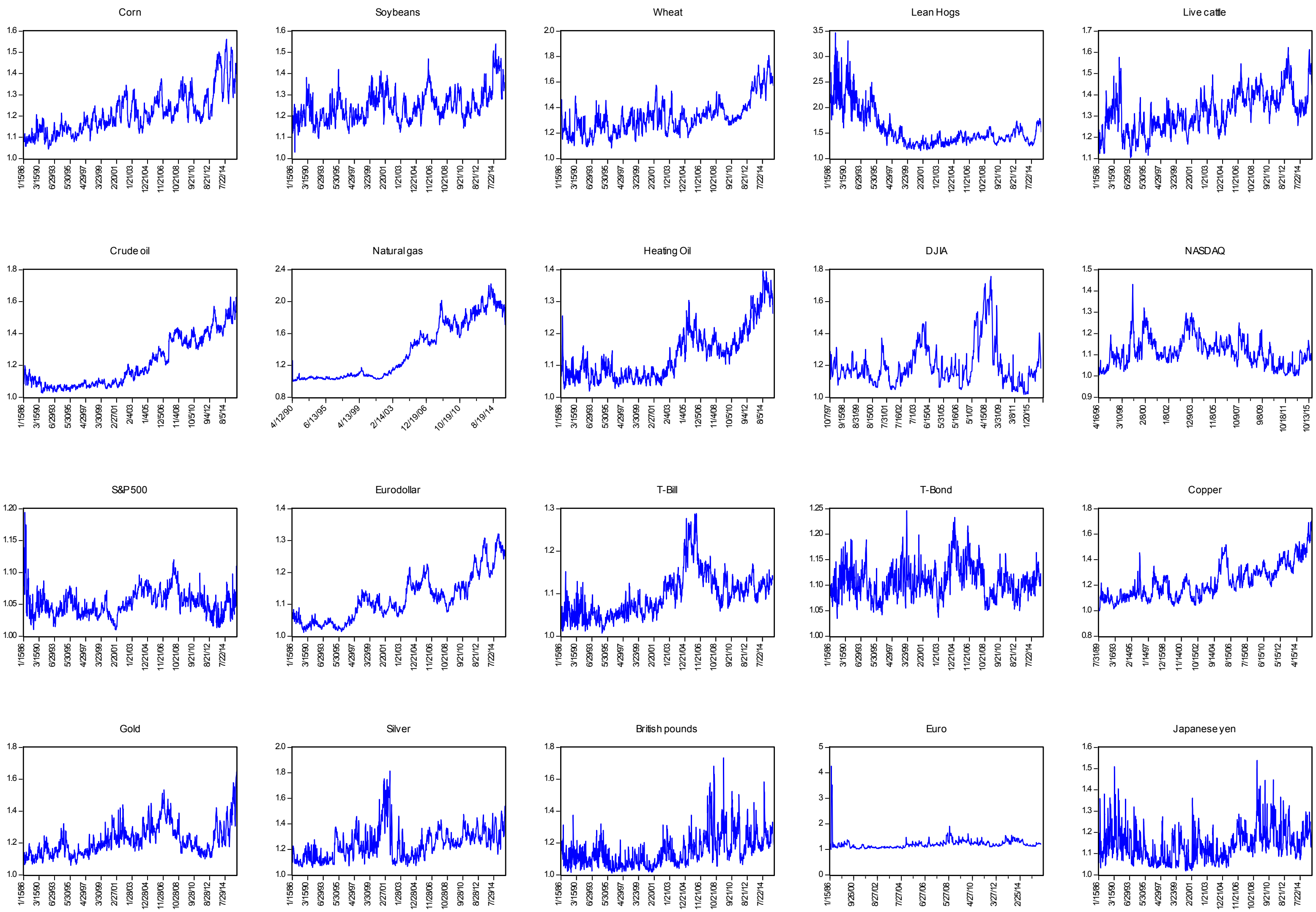


Table 5.3.3 Descriptive Statistics and Results of the ADF Test of Nonstationarity

Categories			Mean	Median	Std. Dev.	Skewness	Kurtosis	J-B	Obs	level of integration
Energy	Crude oil	EV	-3.846	-3.869	0.361	0.788	4.460	264.005	1373	I(0)
		UEV	-4.369	-4.343	0.650	-0.459	4.587	192.387	1373	I(0)
		INDADSP	0.443	0.338	0.279	0.627	1.937	154.702	1374	I(1)
		INDEXSP	0.053	0.023	0.065	1.577	4.578	712.273	1374	I(1)
		T	1.211	1.148	0.156	0.706	2.167	153.923	1374	I(1)
	Natural Gas	EV	-2.407	-2.296	0.842	-1.803	10.605	3751.427	1271	I(0)
		UEV	-2.452	-2.342	0.849	-1.734	10.199	3381.324	1271	I(0)
		INDADSP	0.454	0.308	0.325	0.450	1.565	151.968	1272	I(1)
		INDEXSP	0.366	0.149	0.398	0.698	2.259	132.440	1272	I(1)
		T	1.380	1.194	0.361	0.533	1.761	141.749	1272	I(1)
	Heating oil	EV	-3.850	-3.885	0.339	0.875	4.807	361.892	1373	I(0)
		UEV	-4.366	-4.356	0.643	-0.658	6.254	704.981	1373	I(0)
		INDADSP	0.315	0.293	0.152	0.804	2.754	151.352	1374	I(1)
		INDEXSP	0.024	0.014	0.028	1.696	5.560	1033.532	1374	I(0)
		T	1.134	1.113	0.083	0.974	3.219	219.787	1374	I(0)
Agricultural	Corn	EV	-1.326	-1.376	0.650	0.381	3.011	33.285	1375	I(0)
		UEV	-2.340	-2.244	1.209	-0.628	4.406	203.517	1375	I(0)
		INDADSP	0.535	0.529	0.178	0.282	2.611	26.910	1376	I(1)
		INDEXSP	0.021	0.014	0.021	1.108	3.814	319.533	1376	I(0)
		T	1.213	1.194	0.099	0.973	3.804	253.984	1376	I(0)
	Soybean	EV	-4.274	-4.314	0.317	0.451	3.089	46.964	1375	I(0)
		UEV	-4.780	-4.730	0.639	-0.701	4.813	301.006	1375	I(0)
		INDADSP	0.637	0.643	0.114	-0.020	2.791	2.599	1376	I(0)
		INDEXSP	0.028	0.021	0.026	1.000	3.543	246.181	1376	I(0)
		T	1.249	1.238	0.074	0.795	3.675	170.922	1376	I(0)
	Wheat	EV	-3.900	-3.949	0.307	0.994	4.034	287.488	1374	I(0)
		UEV	-4.382	-4.376	0.610	-0.395	4.939	251.015	1374	I(0)
		INDADSP	0.704	0.698	0.136	0.130	2.773	6.847	1375	I(1)
		INDEXSP	0.071	0.031	0.107	2.373	7.978	2709.687	1375	I(1)
		T	1.328	1.312	0.132	0.964	3.916	261.212	1375	I(0)
Metal	Copper	EV	-4.258	-4.301	0.336	0.868	3.968	212.056	1289	I(0)
		UEV	-4.759	-4.720	0.619	-0.593	4.556	205.656	1289	I(0)
		INDADSP	0.573	0.566	0.192	0.303	2.626	27.292	1290	I(1)
		INDEXSP	0.052	0.030	0.070	2.459	9.345	3464.048	1290	I(0)
		T	1.230	1.206	0.135	0.728	2.997	113.914	1290	I(0)
	Gold	EV	-4.697	-4.700	0.420	0.084	3.224	4.477	1374	I(0)
		UEV	-5.200	-5.159	0.668	-0.615	4.629	238.513	1374	I(0)

		INDADSP	0.678	0.721	0.188	-0.495	2.133	99.308	1375	I(0)
		INDEXSP	0.037	0.015	0.053	1.989	6.306	1532.642	1375	I(1)
		T	1.218	1.196	0.097	1.109	4.312	380.649	1375	I(0)
	Sliver	EV	-4.099	-4.098	0.348	0.280	2.898	18.529	1374	I(0)
		UEV	-4.564	-4.526	0.615	-0.665	4.860	299.488	1374	I(0)
		INDADSP	0.741	0.794	0.159	-1.025	3.498	255.160	1375	I(1)
		INDEXSP	0.047	0.032	0.056	1.969	7.227	1911.910	1375	I(0)
		T	1.244	1.233	0.123	1.088	4.888	475.555	1375	I(0)
Livestock	Lean hogs	EV	-4.227	-4.256	0.215	0.720	3.731	149.419	1375	I(0)
		UEV	-4.783	-4.710	0.592	-1.137	6.044	826.886	1375	I(0)
		INDADSP	0.837	0.844	0.107	-0.504	2.917	58.747	1376	I(0)
		INDEXSP	0.411	0.092	0.585	1.428	3.658	492.360	1376	I(1)
		T	1.576	1.437	0.363	1.720	5.902	1161.444	1376	I(0)
	Live cattle	EV	-4.769	-4.794	0.233	0.544	3.345	74.544	1375	I(0)
		UEV	-5.331	-5.220	0.628	-1.339	6.889	1277.177	1375	I(0)
		INDADSP	0.709	0.678	0.155	0.017	1.850	75.888	1376	I(1)
		INDEXSP	0.050	0.037	0.047	1.169	3.541	330.219	1376	I(0)
		T	1.321	1.321	0.094	0.263	2.841	17.274	1376	I(0)
Foreign Currency	British pounds	EV	-5.230	-5.247	0.277	0.879	4.521	308.930	1372	I(0)
		UEV	-5.759	-5.680	0.579	-0.954	5.564	583.636	1372	I(0)
		INDADSP	0.567	0.562	0.162	0.240	2.324	39.240	1373	I(0)
		INDEXSP	0.016	0.005	0.022	1.645	5.098	870.842	1373	I(0)
		T	1.151	1.130	0.102	1.481	6.263	1110.701	1373	I(0)
	Euro	EV	-5.115	-5.108	0.253	-0.008	3.919	31.512	896	I(0)
		UEV	-5.650	-5.568	0.592	-1.233	6.457	673.008	896	I(0)
		INDADSP	0.629	0.692	0.186	-0.488	1.968	75.428	897	I(0)
		INDEXSP	0.042	0.015	0.102	6.227	46.970	78057.660	897	I(0)
		T	1.209	1.180	0.202	7.837	96.990	339356.000	897	I(0)
	Japanese yen	EV	-5.037	-5.055	0.237	0.807	4.375	257.659	1375	I(0)
		UEV	-5.539	-5.468	0.603	-1.115	7.174	1283.312	1375	I(0)
		INDADSP	0.587	0.559	0.141	0.104	2.540	14.587	1376	I(0)
		INDEXSP	0.006	0.003	0.010	2.698	12.490	6832.414	1376	I(0)
		T	1.143	1.134	0.077	1.019	4.694	402.703	1376	I(0)
Fixed-income	Eurodollar	EV	-3.480	-3.896	1.356	0.639	2.181	132.035	1375	I(1)
		UEV	-3.828	-4.237	1.382	0.591	2.293	108.666	1375	I(1)
		INDADSP	0.276	0.311	0.144	0.096	1.904	70.987	1376	I(1)
		INDEXSP	0.029	0.028	0.021	0.553	2.594	79.436	1376	I(0)
		T	1.119	1.108	0.073	0.671	2.721	107.657	1376	I(0)
	10-Year T-note	EV	-4.362	-4.437	0.461	0.496	2.390	77.693	1375	I(0)
		UEV	-4.866	-4.865	0.685	-0.356	3.859	71.322	1375	I(0)
		INDADSP	0.244	0.253	0.096	0.552	2.856	71.088	1376	I(1)

		INDEXSPT	0.014	0.011	0.016	2.234	8.373	2799.293	1376	I(0)
		INDEXSPT	1.097	1.089	0.052	0.936	3.992	257.581	1376	I(0)
	U.S. T-bond	EV	-3.984	-4.189	0.818	0.363	1.818	110.212	1375	I(1)
		UEV	-4.405	-4.584	0.994	0.139	2.651	11.405	1375	I(0)
		INDADSP	0.255	0.244	0.054	0.882	3.604	199.392	1376	I(1)
		INDEXSP	0.018	0.017	0.013	0.689	3.053	109.070	1376	I(0)
		T	1.108	1.104	0.032	0.663	3.602	121.694	1376	I(0)
Stock index	DJIA	EV	-4.589	-4.656	0.396	0.900	3.922	127.612	749	I(0)
		UEV	-5.150	-5.153	0.649	-0.401	4.613	101.267	749	I(0)
		INDADSP	0.520	0.488	0.170	0.936	3.400	114.477	750	I(1)
		INDEXSP	0.040	0.016	0.057	2.263	8.226	1493.299	750	I(0)
		T	1.200	1.166	0.134	1.617	5.726	558.860	750	I(0)
	NASDAQ	EV	-4.187	-4.268	0.457	0.573	2.604	55.720	911	I(0)
		UEV	-4.778	-4.758	0.702	-0.595	4.856	184.611	911	I(0)
		INDADSP	0.316	0.326	0.102	-0.162	3.369	9.185	912	I(0)
		INDEXSP	0.030	0.016	0.033	1.249	3.752	258.648	912	I(0)
		T	1.113	1.106	0.067	0.699	3.538	85.345	912	I(0)
	S&P500	EV	-4.680	-4.749	0.407	1.104	4.800	465.045	1375	I(0)
		UEV	-5.228	-5.213	0.648	-0.657	6.840	943.781	1375	I(0)
		INDADSP	0.142	0.140	0.039	0.365	2.859	31.637	1376	I(0)
		INDEXSP	0.017	0.012	0.015	0.843	2.916	163.378	1376	I(0)
		T	1.052	1.049	0.021	1.172	6.819	1150.965	1376	I(0)

The results of **Table 5.3.3** indicate that the series of weekly expected and unexpected volatility exhibit different levels of integration than the speculative indices, except for soybean, all foreign currencies, the NASDAQ, and S&P 500 stock indices. These results indicate that the ARDL mode is appropriate to use to capture the long-run and short run relationship between volatilities and the three speculative indices.

Table 5.3.4 ARDL Bounds Test and Estimated Long Run Coefficients Using the Selected ARDL Model for Expected Volatility

Commodities/ Financial Assets (Dependent Variable: EV)		ARDL Bounds test			Estimated Long Run Coefficients			ECM	Diagnostic tests		
		Selected Model	F-statistic	Cointegration	INDADSP	INDEXSP	T	ETC(-1)	RESET	ARCH-LM	Heteroscedasticity
Energy	Crude oil	ARDL(6,2,1,3)	12.811***	Present	-2.515**	0.474	4.413**	-0.062***	0.141(0.869)	1.465(0.111)	2.087(0.009)
	Natural Gas	ARDL(2,3,2,0)	36.870***	Present	-0.252	-0.901	1.548	-0.197***	12.279(0.000)	1.819(0.053)	26.373(0.000)
	Heating oil	ARDL(4,0,1,1)	16.999***	Present	-0.745	2.499	0.205	-0.097***	0.422(0.656)	1.027(0.424)	2.845(0.003)
Agricultural	Corn	ARDL(3,2,5,1)	16.065***	Present	1.167	-4.456	0.071	-0.083***	0.469(0.626)	0.814(0.663)	1.152(0.307)
	Soybean	ARDL(6,1,2,1)	13.159***	Present	0.735	-0.407	0.053	-0.063***	0.764(0.466)	1.387(0.145)	5.051(0.000)
	Wheat	ARDL(6,0,0,2)	14.769***	Present	0.095	-2.325***	2.012**	-0.080***	4.538(0.011)	1.773(0.019)	1.476(0.134)
Metal	Copper	ARDL(2,1,0,1)	7.017***	Present	0.375	0.396	-0.103	-0.029***	3.748(0.024)	1.292(0.229)	6.352(0.000)
	Gold	ARDL(4,4,7,4)	9.222***	Present	1.537***	1.458	-1.349	-0.048***	1.386(0.251)	1.242(0.233)	2.902(0.000)
	Sliver	ARDL(1,2,2,2)	6.117***	Present	0.449	1.128	-0.449	-0.033***	0.494(0.610)	1.446(0.173)	2.129(0.020)
Livestock	Lean hogs	ARDL(2,2,5,2)	13.088***	Present	-0.249	-0.266	0.311	-0.069***	2.939(0.053)	1.385(0.146)	2.888(0.000)
	Live cattle	ARDL(5,1,3,1)	8.073***	Present	-0.213	-0.098	0.696	-0.043***	0.352(0.703)	0.763(0.719)	2.403(0.003)
Foreign Currency	British pounds	ARDL(3,1,5,0)	9.904***	Present	1.186***	0.057	-1.274	-0.034***	1.010(0.365)	0.975(0.479)	2.880(0.001)
	Euro	ARDL(3,0,0,4)	3.565	Non-Present	0.190	-0.141	-1.050	-0.021**	0.292(0.747)	2.866(0.000)	5.938(0.000)
	Japanese yen	ARDL(3,12,2,2)	12.387***	Present	-0.059	4.845	-0.264	-0.065***	0.349(0.706)	0.617(0.863)	2.317(0.001)
Fixed-income	Eurodollar	ARDL(12,0,0,0)	1.644	Non-Present	12.22	-26.557	-7.342	-0.017**	1.763(0.172)	1.177(0.283)	3.298(0.000)
	T-note	ARDL(3,6,1,0)	5.274**	Present	-2.186	-26.318***	13.021***	-0.018**	8.017(0.000)	0.743(0.741)	2.755(0.001)
	U.S. T-bond	ARDL(3,0,10,2)	1.850	Non-Present	6.509	-32.818	-11.632	-0.007*	4.914(0.008)	0.871(0.569)	1.954(0.010)
Stocks index	DJIA	ARDL(2,2,0,0)	10.187***	Present	0.785	3.026	-1.67	-0.078***	1.761(0.173)	0.391(0.981)	0.959(0.459)
	NASDAQ	ARDL(4,5,1,1)	5.674***	Present	-0.768	2.565	1.742	-0.041***	5.262(0.005)	0.778(0.703)	2.385(0.003)
	S&P500	ARDL(9,2,12,9)	8.213***	Present	-4.647**	2.354	10.438**	-0.063***	5.829(0.003)	1.324(0.179)	1.557(0.021)

, **, * statistically significant at the 10%, 5% and 1% level of significance. The null hypothesis is that the series is nonstationary. The null hypothesis for the ARDL bounds test is that no long-run relationship exists between the two variables. The appropriate ARDL model is selected on the basis of the AIC. The nulls for the diagnostic tests are: (ARCH-LM) No serial correlation in the residuals (Heteroscedasticity) No Heteroscedasticity in the error terms (RESET) No functional-form misspecification respectively.*

Table 5.3.4 presents the results for the most appropriate ARDL (p, q) model for the different commodities/financial assets. The dependent variable for the ARDL bounds test is weekly expected volatility and the independent variables are the lagged values of expected volatility, INDADSP, INDEXSP and T. The F-statistic associated with the ARDL bounds test indicates that there is cointegration between the expected volatility and the speculative indices for most commodities/financial assets, except for the Euro and the Eurodollar. **Table 5.3.4** also shows that there is a negative long run relationship between INDADSP and EV for crude oil and the S&P500 stock index at the 5% level of significance and a positive relationship between INDADSP and EV for gold and the British pound at the 1% level of significance. There is a long-term negative relationship between INDEXSP and EV for wheat and T-notes at the 1% level of significance. The long-term relationships between T and EV are positive for crude oil, wheat, T-note, and the S&P500 stock index at the 5% level of significance. The ECM column shows that relationship in the short term, the coefficients of ETC(-1) are all negative and significant at the 10% level of significance, which suggests that there is a short-term relationship between these variables. Natural gas has the highest absolute value of the coefficient of ETC(-1) in Equation (4.10), which means that 19.7% of the disequilibrium between expected volatility and speculation is corrected within one year. The diagnostic tests indicate that all the residuals are free from autocorrelation at the 1% level of significance, except for Euro and natural gas (no autocorrelation at the 10% level of significance). Although the null hypothesis of RESET can be rejected, in other words there is functional-form misspecification, for natural gas, T-note, NASDAQ stock index, and S&P 500 stock index at 1% level of significance and for wheat and copper at 5% level of significance and for lean hogs at 10% level of significance, according to Pesaran et al. (2001), this type of functional-form misspecification may be caused by the presence of non-linear effects or asymmetries in the adjustment process. I do not address this, since both equations **(4.6)** and **(4.7)** pass the most important requirement, which is a lack of serial correlation in the error terms.

Table 5.3.5 ARDL Bound Tests and Estimated Long Run Coefficients Using the Selected ARDL Model for Unexpected Volatility

Commodities/ Financial Assets (Dependent Variable: UEV)		ARDL Bound test			Estimated Long Run Coefficients			ECM	Diagnostic tests		
		Selected Model	F-statistics	Cointegration	INDADSP	INDEXSP	T	ETC(-1)	RESET	ARCH-LM	Heteroscedasticity
Energy	Crude oil	ARDL(11,1,3,2)	14.161***	Present	-2.100**	0.106	3.707*	-0.303***	7.598(0.001)	0.596(0.880)	1.028(0.425)
	Natural Gas	ARDL(2,3,2,1)	37.735***	Present	-0.523	-1.233	2.130	-0.240***	9.878(0.000)	1.317(0.184)	25.670(0.000)
	Heating oil	ARDL(8,0,2,2)	17.340***	Present	-1.031	1.424	0.847	-0.347***	18.594(0.000)	1.466(0.146)	1.775(0.033)
Agricultural	Corn	ARDL(7,0,1,0)	21.881***	Present	1.230	0.914	-1.032	-0.418***	6.672(0.001)	1.352(0.163)	2.367(0.007)
	Soybean	ARDL(8,1,1,4)	19.179***	Present	0.952*	-2.831	-0.34	-0.376***	2.985(0.051)	0.903(0.561)	1.378(0.138)
	Wheat	ARDL(4,1,1,3)	38.572***	Present	0.502	-1.432***	0.805	-0.448***	12.502(0.000)	1.141(0.327)	0.891(0.556)
Metal	Copper	ARDL(12,5,0,4)	7.284***	Present	0.205	0.810	-0.250	-0.226***	1.399(0.247)	0.813(0.663)	1.117(0.316)
	Gold	ARDL(8,0,1,6)	16.830***	Present	1.424***	0.807	-1.361	-0.311***	0.546(0.579)	1.094(0.356)	1.293(0.182)
	Sliver	ARDL(12,7,1,0)	6.798***	Present	0.472	1.441	-0.649	-0.207***	0.250(0.779)	0.632(0.851)	0.614(0.922)
Livestock	Lean hogs	ARDL(9,1,0,1)	22.724***	Present	-0.061	-0.064	-0.109	-0.598***	4.921(0.007)	1.192(0.271)	1.232(0.245)
	Live cattle	ARDL(12,0,0,0)	13.408***	Present	-0.212	-0.991	0.711	-0.480***	0.283(0.7530)	1.094(0.357)	1.574(0.074)
Foreign Currency	British pounds	ARDL(12,4,0,0)	10.448***	Present	-0.013	-4.405*	1.131*	-0.343***	4.582(0.010)	1.151(0.305)	1.699(0.030)
	Euro	ARDL(10,0,0,0)	8.727***	Present	-0.510	-0.510	0.862	-0.365***	1.608(0.201)	0.841(0.632)	0.739(0.725)
	Japanese yen	ARDL(11,0,1,0)	16.736***	Present	-0.064	-2.574	0.585	-0.481***	10.764(0.000)	0.934(0.511)	0.868(0.601)
Fixed-income	Eurodollar	ARDL(12,0,0,0)	1.523	Non-present	11.361	-17.754	-7.415	-0.030*	11.030(0.000)	0.991(0.462)	3.757(0.000)
	T-note	ARDL(12,0,1,0)	6.723***	Present	-2.602	-21.157***	11.976***	-0.173***	1.995(0.136)	1.217(0.252)	0.777(0.713)
	U.S. T-bond	ARDL(9,0,0,1)	2.443	Non-present	-3.999	-32.426	17.801	-0.047*	9.720(0.000)	1.029(0.432)	0.770(0.692)
Stock index	DJIA	ARDL(6,2,0,0)	12.326***	Present	-0.168	6.107**	-2.135*	-0.323***	1.124(0.325)	0.612(0.866)	2.348(0.008)
	NASDAQ	ARDL(10,0,1,0)	7.000***	Present	-1.607	1.312	2.845	-0.236***	8.173(0.000)	1.154(0.303)	0.910(0.548)
	S&P500	ARDL(12,0,3,0)	10.084***	Present	-3.294*	4.256	0.530	-0.283***	21.567(0.000)	0.648(0.837)	0.515(0.953)

, **, * statistically significant at the 10%, 5% and 1% level of significance. The null hypothesis is that the series is nonstationary. The null hypothesis for the ARDL bounds test is that no long-run relationship exists between the two variables. The appropriate ARDL model is selected on the basis of the AIC. The nulls for the diagnostic tests are: (ARCH-LM) No serial correlation in the residuals (Heteroscedasticity) No Heteroscedasticity in the error terms (RESET) No functional-form misspecification respectively. The value in parentheses of diagnostic tests represent p-values.*

Table 5.3.5 Table 5.3.5 indicates the appropriate ARDL (p, q) model for different commodities/financial assets. The dependent variable for ARDL bound test is weekly unexpected volatility of spot returns (UEV) and the independent variables are the lag of expected volatility, INDADSP, INDEXSP and T indices. The F-statistics of ARDL bound test suggests that there is cointegration among the UEV and trading behavior for all contracts, except for Eurodollar contracts, which is the same result with EV and T-bond. For the long-term relationship between INDADSP and UEV, the results are similar with that of INDADSP and EV, besides there is also a positive relationship for soybean at 10% level of significance. For the long-term relationship between INDEXSP and UEV, the results are similar with that of INDEXSP and EV as well, besides, there is a long-term negative relationship for British pounds at 10% level of significance and a positive relationship for DJIA at 5% level of significance. The long-term relationship are positive between T and UEV for Crude oil, British Pounds and T-note and negative for DJIA. In the short term, the ETC(-1) are all negative and significant at 10% level of significant, which suggests that there is short-term relationship among these variables. Moreover, lean hogs has the highest absolute value of the coefficient of ETC(-1) in Equation (4.11), which means that 59.8% of the disequilibrium between unexpected volatility and speculation is corrected within one year. The diagnostics test indicates that all the functions are free from autocorrelation at 1% level of significance. The last but not the least, the percentage of the disequilibrium between unexpected volatility and speculation is corrected more than that between expected volatility and speculation for all commodities/financial assets.

Table 5.3.6 Toda-Yamamoto Causality test Results for Expected Volatility

Commodities/ Financial Assets		INDADSP → EV	INDEXSP → EV	T → EV	EV → INDADSP	EV → INDEXSP	EV → T
Energy	Crude oil	11.606(0.236)	12.762(0.174)	17.761(0.038)**	15.011(0.091)*	15.605(0.076)*	13.997(0.122)
	Natural Gas	1.896(0.388)	2.209(0.331)	4.078(0.130)	0.741(0.691)	2.108(0.349)	0.253(0.881)
	Heating oil	6.764(0.873)	20.9(0.052)*	17.337(0.137)	43.608(0.000)***	12.26(0.425)	14.091(0.295)
Agricultural	Corn	11.58(0.396)	24.424(0.011)**	6.273(0.855)	19.788(0.048)**	10.918(0.450)	22.035(0.024)**
	Soybean	14.517(0.206)	24.511(0.011)**	17.199(0.102)	20.044(0.045)**	22.674(0.020)**	14.357(0.214)
	Wheat	14.659(0.329)	19.067(0.121)	25.917(0.017)**	10.059(0.689)	7.321(0.885)	16.025(0.248)
Metal	Copper	25.808(0.040)**	29.211(0.015)**	34.343(0.003)***	16.008(0.689)	17.266(0.303)	13.574(0.558)
	Gold	4.688(0.096)*	4.453(0.108)	7.168(0.028)**	3.335(0.189)	19.087(0.000)***	10.674(0.005)***
	Sliver	13.563(0.060)*	40.676(0.000)***	25.366(0.001)***	4.246(0.751)	5.577(0.590)	6.586(0.473)
Livestock	Lean hogs	13.934(0.455)	27.257(0.018)**	22.886(0.062)*	26.142(0.025)**	16.515(0.283)	13.501(0.488)
	Live cattle	5.545(0.852)	15.547(0.113)	23.585(0.009)***	21.453(0.018)**	17.002(0.074)*	11.227(0.340)
Foreign Currency	British pounds	23.463(0.005)***	18.437(0.030)**	7.105(0.626)	18.834(0.027)**	9.219(0.417)	7.232(0.613)
	Euro	5.196(0.983)	20.321(0.120)	36.908(0.001)***	16.901(0.262)	32.693(0.003)***	23.625(0.051)*
	Japanese yen	12.213(0.057)*	23.039(0.001)***	15.818(0.015)**	11.562(0.073)*	19.257(0.004)***	14.031(0.029)**
Fixed-income	Eurodollar	11.495(0.320)	13.983(0.174)	10.811(0.373)	11.578(0.314)	15.792(0.106)	11.639(0.310)
	T-note	28.342(0.002)***	14.248(0.162)	12.288(0.266)	12.641(0.245)	3.362(0.972)	10.092(0.433)
	T-bond	14.345(0.279)	34.657(0.001)***	20.239(0.063)*	6.42(0.893)	3.718(0.988)	8.451(0.749)
Stock index	DJIA	5.073(0.407)	1.645(0.896)	3.365(0.644)	10.7(0.058)*	4.252(0.514)	10.649(0.059)*
	NASDAQ	8.543(0.129)	1.085(0.955)	5.333(0.377)	17.956(0.003)***	4.021(0.546)	6.108(0.296)
	S&P500	23.211(0.183)	30.016(0.037)**	33.605(0.014)**	12.02(0.846)	43.69(0.001)***	23.701(0.165)

, **, * statistically significant at the 10%, 5% and 1% level of significance The value in parentheses represent p-values. The lag length selection was based on AIC. → denotes one-way causality.*

Table 5.3.7 Toda-Yamamoto Causality test Results for Unexpected Volatility

Commodities/ Financial Assets		INDADSP → UEV	INDEXSP → UEV	T → UEV	UEV → INDADSP	UEV → INDEXSP	UEV → T
Energy	Crude oil	6.520(0.687)	10.434(0.317)	4.992(0.835)	2.856(0.970)	16.531(0.057)*	7.489(0.586)
	Natural Gas	1.504(0.471)	1.332(0.514)	4.127(0.127)	0.857(0.652)	2.474(0.290)	0.227(0.893)
	Heating oil	10.547(0.649)	11.614(0.560)	6.948(0.905)	20.282(0.088)*	18.058(0.155)	21.484(0.064)*
Agricultural	Corn	6.768(0.562)	7.996(0.434)	7.944(0.439)	9.736(0.284)	4.264(0.833)	10.502(0.232)
	Soybean	7.314(0.503)	17.978(0.021)**	15.103(0.057)*	11.059(0.198)	7.035(0.533)	6.128(0.633)
	Wheat	11.115(0.677)	16.395(0.290)	18.863(0.170)	13.937(0.454)	15.86(0.322)	14.955(0.381)
Metal	Copper	17.41(0.295)	15.14(0.441)	25.402(0.045)**	12.579(0.635)	11.001(0.753)	15.165(0.440)
	Gold	14.007(0.233)	17.137(0.104)	20.202(0.043)**	16.915(0.110)	11.998(0.364)	14.477(0.208)
	Sliver	14.023(0.122)	5.857(0.754)	2.682(0.976)	9.068(0.431)	5.578(0.781)	10.762(0.292)
Livestock	Lean hogs	13.338(0.500)	12.385(0.576)	13.621(0.478)	8.554(0.859)	14.693(0.400)	35.396(0.001)***
	Live cattle	11.48(0.571)	22.132(0.053)*	18.833(0.128)	13.295(0.425)	27.596(0.010)***	10.851(0.623)
Foreign Currency	British pounds	17.43(0.234)	13.758(0.468)	8.005(0.889)	10.315(0.739)	26.105(0.025)**	30.209(0.007)***
	Euro	27.054(0.078)*	16.135(0.583)	19.798(0.344)	22.799(0.199)	12.044(0.845)	35.413(0.008)***
	Japanese yen	2.890(0.823)	5.376(0.497)	8.651(0.194)	2.835(0.829)	7.348(0.290)	4.948(0.550)
Fixed-income	Eurodollar	13.882(0.127)	7.102(0.627)	10.167(0.337)	20.430(0.015)**	7.595(0.575)	15.884(0.069)*
	T-note	10.213(0.422)	8.813(0.550)	10.578(0.391)	13.288(0.208)	5.803(0.832)	13.289(0.208)
	T-bond	8.264(0.764)	24.863(0.016)**	20.367(0.060)*	14.371(0.278)	8.857(0.715)	13.87(0.309)
Stock index	DJIA	10.733(0.03)	1.470(0.832)	6.981(0.137)	3.129(0.536)	3.234(0.519)	1.565(0.815)
	NASDAQ	8.206(0.414)	5.487(0.705)	5.657(0.686)	20.328(0.009)***	9.509(0.301)	7.535(0.480)
	S&P500	19.490(0.362)	12.168(0.839)	13.981(0.730)	6.031(0.996)	28.406(0.056)*	26.166(0.096)*

, **, * statistically significant at the 10%, 5% and 1% level of significance. The value in parentheses represent p-values. The lag length selection was based on AIC. → denotes one-way causality.*

Table 5.3.6 and **Table 5.3.7** provide the results of Toda-Yamamoto causality tests of relationships between expected volatility and speculation and unexpected volatility and speculation respectively. In general, **Table 5.3.6** indicates that the null hypothesis of no causal effect of INDADSP, INDEXSP and T on expected volatility is rejected for 6, 10 and 11, respectively, of 20 commodities/financial assets and the null hypothesis of no causal effect of expected volatility on INDADSP, INDEXSP and T is rejected for 10, 7 and 5, respectively, of 20 commodities/financial assets. **Table 5.3.7** indicates that the null hypothesis of no causal effect of unexpected volatility on INDADSP, INDEXSP and T, is rejected for 3, 4, and 6, respectively, of 20 commodities/financial assets and the null hypothesis of no causal effect of INDADSP, INDEXSP and T on unexpected volatility is rejected for 1, 3, and 4, respectively of 20 commodities/financial assets. In general, the speculative activities are more likely to cause expected volatility than the relationship vice versa. The unexpected volatility is more likely to cause speculative activates than the relationship vice versa. Moreover, the causality relationship between speculative activities and expected volatility are stronger than the relationship between speculative activities and unexpected volatility according to these 20 commodities and financial assets.

In details, there is a lead effect of INDADSP on expected volatility for metal commodities, British pounds, Japanese yen and T-note, and on unexpected volatility for the Euro. There is a lead effect of INDEXSP on expected volatility for heating oil, corn, soybean, copper, silver, lean hogs, the British pounds, Japanese yen, the T-bond, and the S&P500 stock index and on unexpected volatility for soybean, live cattle, and the U.S. T-bond. There is a lead effect of T on expected volatility for crude oil, wheat, all metal, all livestock, the Euro, Japanese yen, the U. S. T-bond, and the S&P500 stock index and on unexpected volatility for soybean, copper, gold, and the U. S. T-bond. There is a lead effect of expected volatility on INDADSP for crude oil, heating oil, corn, soybean, all Livestock, British pounds, Japanese yen, the DJIA, and the NASDAQ stock index, on INDEXSP for crude oil, soybean, gold, live cattle, the Euro, the Japanese yen and S&P500 stock index, and on T for corn, gold, Euro, Japanese yen and DJIA. There is a lead effect of UEV on INDADSP for heating oil, Eurodollar, and NASDAQ stock index, on INDEXSP for crude oil, live cattle, British Pounds, and S&P500 stock index and, finally, on T for heating oil, lean hogs, British Pounds, Euro, Eurodollar, and S&P500 stock index.

There is a bidirectional causality between expected volatility and INDADSP for British pounds, and the Japanese yen, between expected volatility and INDEXSP for soybean, Japanese yen and the S&P500 stock index, and between expected volatility and T for gold, the Euro, and the Japanese yen. There is bidirectional causality between unexpected volatility and INDEXSP for live cattle. There is a lack of bidirectional causality between unexpected volatility and the speculative indices INDADSP and T for all 20 commodities/financial assets.

6. Conclusion

This thesis uses the GJR-GARCH model to model the weekly expected and unexpected volatility of spot returns for 20 commodities/financial assets and the ARDL bounds test and Toda-Yamamoto causality test to investigate the dynamic lead and lag relations between volatilities and speculation.

The results of the first analysis indicate that the expected volatility of spot returns for the 20 commodities/financial assets over the period 1986 to 2015 can be well captured by a GJR-GARCH model along with a t distribution for the error term. The expected and unexpected volatility of spot returns for 20 commodities/financial assets are stationary at levels.

I use Shanker's (2017) weekly estimates of the index of adequate speculation, the index of excess speculation and Working's (1960) speculative index T, as the measures of speculation for the 20 commodities/financial assets over the period 1986-2015. As Shanker (2017) notes, the index of adequate speculation measures the amount of speculation which is just sufficient to equal unbalanced hedging, and the index of excess speculation measures speculation in excess of that required to meet unbalanced hedging. Shanker (2017) establishes that these together correctly estimate Working's (1960) conceptual speculative index which he defines as the ratio of speculation to unbalanced hedging, while Working's (1960) formula for his speculative index T does not.

In this thesis, after matching weekly volatilities with weekly values of the three speculative indices, I establish the series of volatilities have a different order of integration than the speculative indices for all commodities/financial asset. Thus, an autoregressive distributed lag model is used to examine the long term and short term relationship between speculation and both expected and unexpected volatility. The results indicate that, for all 20 commodities/financial assets, there is a statistically significant short term dynamic relationship between both expected and unexpected volatility and speculation. In the long run, the results show that, for the majority of the commodities/financial assets, there is a positive relationship between both expected and unexpected volatility and T index, a negative relationship between both expected and unexpected volatility and INDEXSP and mix relationship between expected and unexpected volatility and INDADSP index.

The results of the Toda-Yamamoto causality test for the 20 commodities/financial assets suggest that the lead-lag relationship depends on the category of the commodities and financial assets, but in general, speculation tends to lead expected volatility more than unexpected volatility for the majority of commodities/financial assets. Expected volatility, rather than unexpected volatility, tend to lead speculation for a majority of commodities/financial assets.

There is a bidirectional causality between expected volatility and INDADSP, INDEXSP, and T for several difference commodities and financial assets. There is also bidirectional causality between unexpected volatility and INDEXSP for live cattle. However, there is no bidirectional causality between unexpected volatility and the speculative indices INDADSP and T for all 20 commodities/financial assets.

The empirical analysis conducted in this thesis could be further extended to account for the effect of structural breaks, if any, in the different series addressed. Since in 2007 an additional category of the aggregate futures and options positions for index traders was added in COT.

7. References

Adrangi, B. and Chatrath, A. (1998). "Money and the Business Cycle: Another Look", *Review of Economics and Statistics*, 70, 680-85.

Alimi, S. R. and Ofonyelu, C. C. (2013). "Toda-Yamamoto Causality Test between Money Market Interest Rate and Expected Inflation: The Fisher Hypothesis Revisited", *European Scientific Journal*, Vol.9, No.7.

Alimi, S. R. (2014). "ARDL Bounds Testing Approach to Cointegration: A Re-examination of Augmented Fisher Hypothesis in an Open Economy", *Asian Journal of Economic Modelling,* 2(2): 103-114.

Anguibi, C. F. C. (2015). "An Investigation of the Long-run and Causal Relationships between Economy. Performance, Investment and Port Sector Productivity in Cote d'lvoire", *Open Journal of Social Sciences*, 3, 29-38.

Antoniou, A. and Holmes, P. (1995). "Future trading, information and spot price volatility: evidence for the FTSE-100 Stock Index Futures contract using GARCH", *Journal of Banking & Finance*, 19, 177 – 129.

Arize, A. C., and Malindretos, J. (1998). "The long-run and short-run effects of exchange-rate volatility on exports: The case of Australia and New Zealand." *Journal of Economics and Finance*, 22(2-3), 43-56.

Basu, D. and J. Miffre. (2009). "Capturing the Risk Premium of Commodity Futures: The Role of Hedging Pressure". *Journal of Banking and Finance,* 37:2652-2664

Bessembinder, H. and Seguin, P. J. (1992). "Futures-trading activity and stock price volatility", *Journal of Finance*, 47, 2015–2034.

Bessembinder, H. and Seguin, P. J. (1993) "Price volatility, trading volume, and market depth: Evidence from futures markets", *Journal of Financial and Quantitative Analysis*, 28, 2015–2034.

Black and Fisher. (1976), "The pricing of commodity contracts", *Journal of Financial Economics*, 3, 167-179.

Bollerslev, T. (1986). "Generalized autoregressive conditional heteroscedasticity", *Journal of Econometrics*, 31, 307–327.

Box, G. E. P. and D. A. Pierce. (1970) "Distribution of residual autocorrelations in autoregressive-integrated moving average time series models", *J. Amer. Statist. Assoc*. 65, 1509-1526.

Bryant, H.L., D.A. Bessler, and M.S. Haigh. (2006). "Causality in Futures Markets", *The Journal of Futures Markets,* 26, 1039–1057.

CFTC. (June 21, 2006). Commission Actions in Response to the Comprehensive Review of the Commitments of Traders Reporting Program. Washington, DC, CFTC.

Chang, E., Chou, R.Y., and Nelling, E.F. (2000). “Market volatility and the demand for hedging in Stock Index Futures”, *The Journal of Futures Markets*, 20, 105–125.

Chatrath, A., Ramchander, S., and Song, F. (1996). “The role of futures trading activity in exchange rate volatility”, *The Journal of Futures Markets*, 16, 561-584.

Chatrath A, Song F, and Adrangi B. (2003). “Futures trading activity and stock price volatility: some extensions”, *Applied Financial Economics,* 13, 655–664.

Chatrath, A., Y. Liang, and F. Song. (1997). "Commitment of Traders, Basis Behavior, and the Issue of Risk Premia in Futures Markets", *The Journal of Futures Markets*, 17, 707–731.

Clifton, E. V. (1985). “The currency futures market and interbank foreign exchange trading”, *The Journal of Futures Markets*, 5, 375-384.

Cox, C. C. (1976). “Futures trading and market information”, *Journal of Political Economy*, 84, 1215-1237.

Danthine, J. (1978). “Information, futures prices, and stabilizing speculation”, *Journal of Economic Theory*, 17, 79-98.

De Roon, F.A., T.E. Nijman, and C. Veld. (2000). "Hedging Pressure Effects in Futures Markets." *The Journal of Finance*, 55, 1437–1456.

Du, X. D., Yu, C. L., and Hayes, D. J. (2011). “Speculation and volatility spillover in the crude oil and agricultural commodity market: A Bayesian analysis”, *Energy Economics,* 33, 497-503.

Dufour, J-M., Pelletier, D., and Renault, E. (2006). “Short run and Long run Causality in Time Series: Inference”, *Journal of Econometrics*, 132, 337-362.

Engle, R. F. (1982). “Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation”, *Econometrica,* 50, 987–1007.

Engle, R. F. and Ng, V.K. (1993). “Measuring and Testing the Impact of News on Volatility”, *The Journal of Finance*, Vol. 48, No. 5., 1749–1778.

Fackler, P. L. and Y. Tian. (1999). “Volatility Models for Commodity Markets.” Proceedings of the NCR-134 Conference on Applied Commodity Price Analysis, Forecasting, and Market Risk Management. Chicago.

French, K., G. W. Schwert., and R. Stambaugh. (1987). “Expected stock returns and volatility”, *Journal of Financial Economics,* 19, 3-29.

Froot, K. A. and Perold, A. F. (1995). “New trading practices and short-run market efficiency”, *Journal of Futures Markets, 15, 731-765..*

Gilbert, C.L. (2010a). "How to Understand High Food Prices", *Journal of Agricultural Economics*, 61, 398–425.

Gilbert, C.L. (2010b). "Speculative Influences on Commodity Futures Prices, 2006-2008", UNCTAD Discussion Papers 197.

Glosten, L. R., Jagannathan, R., and Runkle, D. E. (1993). “On the Relation between the Expected Value and the Volatility of the Nominal Excess Returns on Stocks”, *Journal of Finance*, 48(5), 1779-1801.

Granger, C. W. J. (1969). “Investigating Causal Relations by Econometric Models and Cross-Spectral Methods”, *Econometrica*, 37, 423–38.

Granger, C. W. J. and Newbold. P. (1974) “Forecasting economic time series” Academic Press, New York.

Grammatikos, T. and Saunders, A. (1986). “Futures price variability: a test of maturity and volume effects”, *Journal of Business,* 59, 319-330.

Gulen, H. and Mayhew, S. (2000). “Stock index futures trading and volatility in international equity markets”, *The Journal of Futures Markets*, 20, 661–685.

Gujarati, D. N. (2006). Essential of Econometrics, 3rd Edition, *McGraw Hill*.

Hicks, John R. (1939). “Value and Capital”, Oxford University Press, Cambridge, U.K.

Irwin, S. H. and D. R. Sanders. (2012b). "Financialization and Structural Change in Commodity Futures Markets", *Journal of Agricultural and Applied Economics*, 44, 371–396.

Ibrahim, M. H. (2015). “Oil and Food Prices in Malaysia: a Nonlinear ARDL Analysis”, *Agricultural and Food Economics a Spring Open Journal*, 3:2.

Johansen, S. and K. Juselius. (1990). "Maximum likelihood estimation and inference on cointegration-with applications to the demand for money", *Oxford Bulletin of Economics and Statistics*, 52, 169-210.

Kalman, R. E. (1960). “A new approach to linear filtering and prediction problem”, *J. Basic Eng.*, Trans. ASME Ser. D82, 35-45.

Keynes and John Maynard. (1930). “A Treatise on Money”, Vol. II (Macmillan, London).

Kyle, A. S. (1985). “Continuous auctions and insider trading”, *Econometrica*, 53, 1315-1335.

Lee, C. L., Stevenson, S., and Lee, M. L. (2014). “Futures Trading, Spot Price Volatility and Market Efficiency: Evidence from European Real Estate Securities Futures”, *Journal of Real Estate Financial Economics*, 48, 299-322.

Lehecka, G. V. (2013). “Hedging and Speculative Pressures: An Investigation of the Relationships among Trading Positions and Prices in Commodity Futures Markets”, Proceedings of the NCCC-134 Conference on Applied Commodity Price Analysis, Forecasting, and Market Risk Management. St. Louis, MO.

Ljung, G.M. and G.E.P. Box. (1978). “On a measure of lack of fit in time series models”, *Biometrika,* 65, 297-303.

McCarthy, J. and Najand, M. (1993). “State space modeling of price and volume dependence: evidence from currency futures”, *The Journal of Futures Markets*, 13, 335-344.

Nelson, D. (1991). "Conditional heteroscedasticity in asset returns: A new approach", *Econometrica*, 59, 347–370.

Pagan, A. and Schwert, G. W. (1990). "Alternative models for conditional stock volatility", *Journal of Econometrics*, 45, 267–290.

Pashigian, B. P. (1986). "The Political Economy of Futures Market Regulation", *Journal of Business*, 59, 55-84.

Pesaran, M. H., Shin, Y. and Smith, R. J. (2001). "Bounds testing approaches to the analysis of level relationships". *Journal of Applied Econometrics*, 16, 289-326.

Poon, S-H. and Granger, C. (2005). "Practical Issues in Forecasting Volatility", *Financial Analysts Journal*, 61, Number 1.

Rambaldi, A. and Doran, H. (1996). "Testing for Granger Non-Causality in Cointegrated Systems Made Easy", Working Papers in *Econometrics and Applied Statistics*, The Department of Econometrics, University of New England, 88

Sanders, D.R., K. Boris, and M. Manfredo. (2004). "Hedgers, Funds, and Small Speculators in the Energy Futures Markets: An Analysis of the CFTC's Commitments of Traders Reports", *Energy Economics*, 26, 425–445.

Sanders, D.R., S.H. Irwin, and R.P. Merrin. (2009). "Smart Money: The Forecasting Ability of CFTC Large Traders in Agricultural Futures Markets", *Journal of Agricultural and Resource Economics*, 34, 276–296.

Sargent, T. J., (1989). "Two models of measurements and the investment accelerator". *J. Polit. Econ.,* 97, 251-287.

Sarwar, G. J. (1987). "Has programmed trading made stock prices more volatile?", *Federal Reserve Bank of St. Louis Review*, 18-29.

Sin, C. Y. and White, H. (1996). "Information Criteria for Selecting Possibly Misspecified Parametric Models", *Journal of Econometrics*, 71(1-2), 207-225.

Singleton, K.J. (2011). "Investor Flows and the 2008 Boom/Bust in Oil Prices", Working paper. Graduate School of Business, Stanford University.

Shanker, L., (2017). "New indices of adequate and excess speculation and their relationship with volatility in the crude oil futures market", *Journal of Commodity Markets* 5, 18-35.

Shrestha, M. B. and K. Chowdhury. (2005). "A Sequential Procedure for Testing Unit Roots in the Presence of Structural Break in Time Series Data: An Application to Nepalese Quarterly Data 1970-2003", *International Journal of Applied Econometrics and Quantitative Studies 2.*

Toda, H. Y. and P. C. B. Phillips. (1994). "Vector autoregressions and causality: a theoretical overview and simulation study", *Econometric Reviews*, 13, 259-285.

Toda, H. Y. and T. Yamamoto (1995). "Statistical Inference in Vector Autoregressions with Possibly Integrated Processes", *Journal of Econometrics*, 66, 225–250.

Umar, M. and Dahalan, J. (2016). “An Application of Asymmetric Toda-Yamamoto Causality on Exchange Rate-inflation Differentials in Emerging Economies”, *International Journal of Economics and Financial Issues*, 6(2), 420-426.

Wang, C. (2003). "The Behavior and Performance of Major Types of Futures Traders", *The Journal of Futures Markets,* 23, 1–31.

Wang, C. (2001). “Investor Sentiment and Return Predictability in Agricultural Futures Markets”, *Journal of Futures Markets*, 21,929–952.

Weaver, R.D. and A. Banerjee. (1990). “Does Futures Trading Destabilize Cash Prices? Evidence for U.S. Live Beef Cattle”, *Journal of Futures Markets*, 10, 41-60.

Working, H. (1960). "Speculation on Hedging Markets", *Food Research Institute Studies*, 1, 185–220.

Yang, J., Balyeat, R. B., and Leatham, D. J. (2005). “Futures Trading Activity and Commodity Cash Price Volatility”, *Journal of Business Finance and Accounting*, 32, 297-323.