

MULTIPLE REGRESSION MODEL FOR COTTON PRICE RETURNS
Analysis the impact of weather, oil price return, and China's economy

Pham Cong Son

International Business
Bachelor's Thesis
Supervisor: Dr. Bruce Hearn
Date of approval: 9 April 2018

Aalto University
School of Business
Bachelor's Program in International Business
Mikkeli Campus

MULTIPLE REGRESSION MODEL FOR COTTON PRICE RETURNS
Analysis the impact of weather, oil price return, and China's economy

Pham Cong Son

International Business
Bachelor's Thesis
Supervisor: Dr. Bruce Hearn
Date of approval: 9 April 2018

Aalto University
School of Business
Bachelor's Program in International Business
Mikkeli Campus

Author: Pham Cong Son
Title of thesis: Multiple regression of cotton price return: Analysis of impact of weather, oil price, and China's economy
Date: 13 April 2018
Degree: Bachelor of Science in Economics and Business Administration
Supervisor: Dr. Bruce Hearn
Objectives The study aims at determining the relationship between cotton price return and oil price return, China's economy, and weather condition, particularly the monsoon season. Furthermore, the study attempts to conduct a multiple regression model to estimate the cotton price return based on oil price return, difference of precipitation in China, India, USA, China's interest rate, China's import, and monsoon.
Summary A multiple regression model is conducted for 263 samples of cotton spot price and independent variables: oil spot price, China's import, China's interest rate, precipitation in USA, China, India, and monsoon period, which are all recorded as monthly data from February of 1994 to December of 2015. The assumption pre-tests for multiple regression model are conducted. Based on the assumption test result, the input set of data can be considered valid for conducting multiple regression model.
Conclusions The empirical results reaffirm the negative relationship between cotton price return and two other variables: China's interest rate, and the change in China's import level and positive relationship between cotton price return and oil price return. The model offers inconclusive conclusion for the relationship of cotton price return and monsoon period.
Key words: <i>cotton price, cotton return, multiple regression model</i>
Language: English
Grade:

COVER PAGE	1
TITLE PAGE.....	2
ABSTRACT	3
TABLE OF CONTENT.....	4
1. INTRODUCTION	6
2. LITERATURE REVIEW.....	7
2.1 Past studies on cotton price prediction.....	7
2.2 Influences on global cotton price.....	9
2.2.1 Weather factors	9
2.2.2 Oil price.....	10
2.2.3 Macroeconomic factors.....	12
2.3 Summary.....	16
2.4 Conceptual Framework	16
3. METHODOLOGY	17
3.1 Theoretical model	17
3.2 Assumption of model.....	18
3.3 Empirical model	19
3.4 Data and Variable selection	20
3.5 Suggested empirical equation	21
3.5.1 Natural logarithm difference.....	21
3.5.2 Percentage difference.....	22
4. FINDINGS	23
4.1 Descriptive statistics.....	23

4.1.1 Natural log difference.....	23
4.1.2 Percentage difference.....	24
4.2 Normality of error test.....	25
4.2.1 Natural log difference.....	25
4.2.2 Percentage difference.....	26
4.3 Linearity test.....	27
4.3.1 Natural log difference.....	27
4.3.2 Percentage difference.....	27
4.4 Correlation test – multicollinearity	28
4.4.1 Natural log difference.....	28
4.4.2 Percentage difference.....	30
4.5 Homoscedasticity test	31
4.5.1 Natural log difference.....	31
4.5.2 Percentage difference.....	32
5. DISCUSSION AND ANALYSIS.....	33
5.1 Natural log difference.....	33
5.2 Percentage difference.....	36
6. CONCLUSIONS.....	37
6.1 Main Findings.....	37
6.2 Limitation of the research.....	38
6.3 Suggestions for Further Research	39
REFERENCE LIST.....	39

1. INTRODUCTION

Agricultural commodity prices are notoriously challenging to predict owing to unprecedented incidents from weather alternation all over the world, and government policies in the marketplace, to changing tastes and technology. Furthermore, owing to the globalization of market, data accessibility, and relatively low reserve stocks of commodities, agricultural markets are driven more significantly and rapidly by diverse information sources than that in the past. Nonetheless, numerous scholars and practitioners across public and private agencies are committed to the research of macroeconomic factors and financial variables that facilitate commodity price estimation. These forecasts are expected to provide useful information to the market and to facilitate price discovery and efficient asset allocation (Vogel & Bange, 1999). Most researches on agricultural commodity price estimation are aimed at general commodities or specific analyses of the most popular ones as crude oil, gold, and wheat. By comparison, there are limited number of studies on cotton price prediction and its main determinants. Previous studies focused on the application of cotton futures pricing and quality-attribute pricing, yet rarely discussed the cotton price influences in the greater scope (Ethridge and Davis, 1982; Cho, 2006; Chen et al., 2014; Isengildina-Massa and MacDonald, 2015).

The cotton industry has experienced an upward volatile trend since 1990 thus far. The global cotton production volume in 2017 equalled to 120862000 bales, which is 38.65% higher than that in 1990 (stata.com, n.d.). Over the period, the top three cotton exporting countries are China, USA, and India while the largest cotton importing one is China (ibid). The price of global cotton suffered from two considerable fluctuation periods due to the spill-over effect of oil price from 2007 to 2009 and China's cotton stockpiling policies from 2011 to 2013 (Nazlioglu and Soytaş 2012; Ji and Yan, 2012; Alvins, 2015). Two events increased volatility of cotton price and enhanced the power of oil price and China's related economic policies over cotton price. Due to the fluctuating state of cotton price over the years, hedgers, including farmers, manufacturers, importers and exporters, have been more concerned about novel methods and studies to improve accuracy in cotton price prediction. Accordingly, the data employed in forecasting methods, especially those variables recently impacting cotton price, are inclined to be examined. Supply and demand factors are claimed to be the principal contributors to the formation and the

movement of the cotton market. In more details, the supply level can include weather factors, oil price while the demand level is influenced to a degree by China's economic policies.

With an aim to investigate the level of influence of such factors discussed before over cotton price, this thesis aspires to answer the following research question:

Which type of regression model is pertinent for evaluating the impact level of weather variables, oil price, China's interest rate, and China's import on cotton price?

Acknowledging that research question, this study will first revisit existing literature on past studies of cotton pricing models, and prior analyses on the impact of weather factors, the global oil price, and China's economy on the global cotton price. Subsequently, the explanation of methodology, data selection, and proposed empirical model are presented. Afterward, the results of assumption tests are examined, followed by the display of the estimated results and further analysis. Finally, the paper summarizes with main findings, limitations of the research, and suggestions for further investigations.

2. LITERATURE REVIEW

2.1 Past studies on cotton price prediction

In past studies of cotton commodity price, the majority of price prediction is based on cotton futures. Ashby (1929) and Brosen et al. (1984) proposed that there is a substantial connection between futures price and commodity spot prices. Many research studies assert that cotton price is discovered in cotton futures price (Hudson et al., 1996; Chavez, 2007; Bhanumurthy et al., 2012). Moreover, it is testified that short-time futures are better forecast of the spot price at maturity than long-time futures (Ashby, 1929) and a sudden change in futures price is forecast to influence spot price for approximately five months at maximum with a lag of one month (Bhanumurthy et al., 2012). However, complete reliance on cotton futures in order to predict the future cotton price can be conducive to big failures. According to Bernake (2008), when futures markets quotes underestimated the level of commodity price gains, this can result in corresponding underpredictions of

overall inflation. Therefore, a new model is necessary to accurately forecast the commodity price leveraging other information such as supply and demand apart from futures price (ibid).

Moreover, previous analyses attempted to forecast cotton price based on its quality attributes, such as the average micronaire reading for the lot, and the lot size in number (Ethrige and Davis, 1982). Nevertheless, this method has a number of limitations. The major assumption of the study is that the valuation of goods is based on their utility attributes, which are the primary driving sources for movement of goods' prices. Moreover, all macroeconomic factors such as interest rate, exchange rate that affect the price and vary with time are omitted. It is indicated in the study that in a specific given year, the quality is deemed to alter price, the level of which is contingent upon the available quality characteristics of goods in that year's crop, and upon the demand for the characteristics. In the international cotton market, the attributes present in each region can be different and the demand will adjust according to regions, therefore variables need to be adjusted to each region and each time periods. In comparison, Cho's analysis expands the scope of study and add weather variables to the model, improving the cotton spot price prediction based on supply factors (quality and weather) (2006). Both studies by Cho (2006) and Ethrige and Davis (1982) applied hedonic regression model, an alternative form of ordinary least squares (OLS) multiple regression model, to estimate the cotton price based on categorical variables. Though both two analyses are tied to cotton quality attribute by each state in US, they herald future feasible application of OLS multiple regression on cotton price in global scale.

The most recent study on cotton price forecasting in conducted by Isengildina-Massa and MacDonald (2015). The study analyzes structural changes in the United States' cotton industry and generate a statistical model that reflects the current drivers of U.S. upland cotton prices. The model took into consideration supply and demand factors such as the cotton supply of United States (U.S.), weather, value of U.S. dollar and changes in China's net imports as a share of world consumption. Although Isengildina-Massa and MacDonald's approach is only applicable within the U.S. market and omits the influence of other commodity prices such as oil price, indicating to some extent the impact of

currency value and cotton imports of china on cotton price prediction. Therefore, the relationship between cotton price and a variety of macro-variables including weather factors, oil price, and Chinese economic activities should be examined.

2.2 Influences on global cotton price

2.2.1 Weather factors

One of the primary reasons leading to the high level of volatility in cotton price is weather shock (Cho, 2006; Isengildina-Massa & MacDonald, 2015). In this comprehensive study of Bhanumurthy et al. (2012), weather shock, a type of exogenous shock, can significantly impact agricultural commodity products, and any alternation in expected weather condition can adversely influence the commodity price in the short term. In respect of cotton production, variation in climatic aspects such as temperature, radiation, and water has been well documented to have substantial impact on cotton yields, in both quality and quantity (Bange et al., 2010; Reddy et al., 1995; Reddy et al., 2004; Cho, 2006).

Due to its salt content, high temperature, and radiation tolerance (Chapagain et al, 2006; Constable and Bange, 2015), cotton is predominantly cultivated in arid and semi-arid yet irrigated regions of the world (Roth et al., 2013) such as China, India, and United States (statista.com, n.d.). Those three countries have generated in total of relatively 63% of world cotton production from 2013/2014 season to 2016/2017 season, posing a substantial impact on the global cotton production and price (Cotton Incorporate, 2018). In such regions, the level of water availability is one of the major determinants of improvement of cotton production. For instance, the length of the rainy season is proven to be proportional to the cotton yield (Blanc et al., 2008, Marteau et al., 2011; Traore et al., 2014) and limited water resources can result in fruit loss and lower cotton output (Bange et al., 2010; Williams et al., 2015). Therefore, it is anticipated that cotton may compete for sparse water resources with other industry and an increase in cotton production can correspondingly aggravate the water pressure significantly in regions which have already faced water problems (Chapagain et al., 2006). This can be exemplified in the case of Aral sea in Uzbekistan, which is the seventh largest cotton producer in 2016 (statista.com, n.d.). Due to improper cropping and irrigation practices,

the expanded cultivation of cotton without control resulted in shrinkage of the Aral sea, widespread salination, and soil depletion in the area.

Despite different levels of rainfall in each country, they are all highly influenced by the monsoon season. Monsoon is commonly described to be seasonal changes in atmospheric circulation and precipitation due to the asymmetric temperature of land and sea, which alters the temperature and precipitation of the corresponding regions. (Trenberth and Stepaniak and Caron, 2000; Zuidema and Fairall, 2007). Monsoon seasons in USA, China, and India arrive during the end of May or beginning of June and last for around 4 months annually (Belles, 2017; Krishnamurti, 2015). During the period, the rainfall level increases and supports the cotton production. On the downside, due to excessive water resources, monsoon often brings about flood and city damage to the nearby residents. Furthermore, the rain could inundate the crop fields and drift away the crops, including cotton. Overall, the relationship between cotton price and weather variables including rainfall level of China, India, and USA, and monsoon period requires further investigation, consequently drive to examine the following hypothesis:

Hypothesis 1: Cotton price returns are negative associated with monsoon (rainy) conditions

2.2.2 Oil price

Firstly, the association between oil price and agricultural commodity price has been under disputed by scholars for many years. Some asserted that there is no link between oil prices and agricultural commodities due to the nature of some agricultural goods, different time episodes or the economies of their producing countries (Yu et al., 2006; Zhang and Reed, 2008). However, recent studies claim that oil price has a substantial power on agricultural commodity prices. Firstly, oil price changes affect agricultural commodity price not only directly through production costs but also via the exchange rate and foreign borrowing (Hanson et al, 1993). Additionally, due to the increasing consumption of biofuel, a replacement for fossil fuels, and hedge strategies against inflation induced by high oil prices, oil market and agricultural market have been experiencing a stronger

correlation (Chang and Su, 2010). This is exemplified by the study undertaken by Saghalian (2010), which proves the cointegration relationships between crude oil and agricultural commodity prices such as corn and soybean. Moreover, during the financial crisis from 2007 to 2009 and the afterward period, the influence of oil price was even more significant (Nazlioglu and Soytaş, 2012). The general explanation is that the sudden increase in oil prices will result in the temporary raise in inflation rate. Since oil is the essential input for equipment processing raw materials and transportation delivering the final output to the customers, the cost of production will also rise (Ji and Fan, 2012; Aziz, 2013). While agricultural commodity markets experienced significant volatility spillover effects by oil market during crisis, the effect lingered and strengthened the correlation (Ji and Yan, 2012). Overall, the volatility of crude oil prices is expected to always influence other commodity prices to some extent, and the relationship is enhanced after crisis (ibid).

While numerous studies asserted the definite connection between oil and agricultural markets, the magnitude of correlation between crude oil price and cotton price has been under dispute recently. Whereas several studies examined the link to be loose (Fadiga and Misra 2007; Plastina, 2010), others advocate the remarkable connection between oil and cotton prices (Baffe, 2007; Mutuc et al, 2010; Silvennoinen and Thorp, 2015). According to Silvennoinen and Thorp (2015), as cotton is moderately linked to biofuel production, it exhibits higher correlation and sensitivity to oil than other commodities. Furthermore, cotton has the largest per-acre energy costs of all agricultural commodities, making cotton prices more vulnerable to oil price changes (Sainsbury, 2015). Besides directly affecting the cotton production cost, oil price can secondarily influence the cotton price owing to substitutability with polyester in textile sector. Polyester fiber industry requires inputs as chemical derivatives of crude oil. Indeed, an incremental alternation in oil price can raise both the production cost and material cost of polyester fiber production. As the result, cotton price will be less exorbitant and its demand will respond more positively. Based on Baffe's thorough analysis on the simultaneous link between oil spot prices and the spot prices of other commodities from 1960 to 2005, monthly adjustments in cotton price were primarily traced back to the unprecedented upward trend in global oil demand, with the elasticity of cotton price to oil price of 14% (2005). Overall, crude oil price has a clearly positive impact on cotton spot price, especially after crisis such as

financial depression between 2007 and 2009. Therefore, the following hypothesis is proposed:

Hypothesis 2: Cotton price returns are positively associated with oil price returns

2.2.3 Macroeconomic factors

The relationship between commodity price and macroeconomic determinants has been examined in previous studies (Chen et al, 2008; Gargano and Timmermann, 2014). The future movements of commodity spot prices are likely to be altered by time-varying storage costs and convenience yields (Brooks et al, 2013; Gargano and Timmermann, 2014). Both of the variables are driven by the discrepancy between supply and demand level of those commodities, which can be manipulated by the current state of overall economy (Pindyck and Rotemberg, 1990; Gargano and Timmermann, 2014). In the extensive research of the power of macroeconomic uncertainty on volatility of commodity price, Joets et al discovered high sensitivity level of agricultural and industrial markets to the variability and the level of macroeconomic uncertainty (2017). Moreover, the 2007-2009 financial crisis is proven to engender an unexpected period of high uncertainty in the price of numerous primary products, particularly those are connected to global economic activity (ibid). As a primary product, cotton is inclined to supply and demand inelasticity, referring that a slight change in demand and supply may lead to greater adjustment in price (Cho, 2006; Chavez, 2007). Hence, supply and demand data, such as export and import, can have a substantial impact on cotton price.

2.2.3.1 Exchange rate

Previous studies have demonstrated the substantial link between world commodity prices and currency values of their major exporting countries (Cashin et al 2004; Engel and West, 2005; Cheung et al., 2005; Chen et al., 2008). The cultivation for agricultural commodities is usually focused on several countries or regions, resulting in the high level of correlation of commodity production with the economy of those countries.

Researchers have proposed two major explanations for the relationship. In the first proposal, it is suggested that change in the commodity price will bring about an adjustment in the corresponding commodity currency. This idea is justified in the study by Chen et al (2014), which testifies the impact of commodity prices over prediction of inflation in small commodity-exporting nations. In all investigated countries, the majority of their export sources are natural and raw products such as oil and natural gas. Consequently, commodity price is rarely just a financial asset but a fundamental determinant to the economies of those countries. In addition, there have been co-movements in prices of the majority of exported commodities in those examined countries, therefore a peak in commodity prices are expected to place growing pressure on the demand for its corresponding currency, resulting in a currency appreciation. However, the method of analysis contains a number of limitations. Firstly, exchange rate is normally difficult to predict, which leads to a mismatch between economic models for evaluating exchange-rate and empirical studies (Isard,1995; Obstfeld and Rogoff, 1996; Mark, 2001; Kilian and Taylor, 2003). Furthermore, the study focuses on small economies which export margins contribute modestly to the global demand and supply, therefore this method does contain potential measurement error in the instance of large economies where simultaneity occurs between global commodity price and their economies, or where the correlation among exported commodities is modest. However, the research strengthens the correlation between commodity returns and it export countries' economies.

In the second theory, exchange rate is considered as a proxy for economic fundamentals and one of the main determinants in forecasting commodity prices (Engel and West, 2005; Cheung et al., 2005; Chen et al., 2008). In order to value the exchange rate of one currency, all the economic fundamentals related to that currency is inquired, including the commodity prices, therefore the exchange rate plays a significant role in determining commodity returns (Zhang et al, 2016). Additionally, as for commodities which are inversely correlated to the real exchange rate shock earn, the expected gain in equilibrium is much higher (ibid).

2.2.3.2 China's economy

China has become a predominant player in both physical and financial commodity markets, particularly agricultural market. Foremost, the country has been the leading importer and exporter of a wide range of commodities ranging from agricultural, energy to metal markets for many years. According to Food and Agriculture Organization of the United Nations (2013), the country accounted for the highest percentage in the global supply of 11 out of 17 listed agricultural commodities (including cotton) in 2010. Moreover, it was the world's second biggest consumer of agricultural commodities after the US in 2009 (USITC, 2011). As regards financial commodity markets, China developed its futures market dramatically, accounting 42.84% of global globally traded contracts in 2012 (WFE, 2013). With the increasing contribution to worldwide agricultural commodity markets, China is recognized as a forerunner in the pricing of globally traded commodities. In the thorough work undertaken by Klotz et al (2014), the relation between Chinese economy with a variety of commodity prices is extensively examined. The authors employed monthly data from 1998 to 2012 on the industrial production and the real interest rate as indicator for Chinese economy and its monetary policy respectively. It is deduced from the study's result that agricultural market has generated a growing interrelationship with the macroeconomic developments in China (ibid). Changes in economic and monetary policies strike China's domestic supply and demand patterns, which consequently cast a spillover effect on specific commodity prices. In addition, agricultural commodity price reacts negatively to an abrupt upward trend in the real interest rate of China at 5% significance, displaying an inverse correlation between agricultural price and real interest rate (ibid:7). The phenomenon is in alignment with the commodity overshooting model by Frankel (2008) and similar empirical study with interest rate in the U.S. (Akram, 2009). While the real interest rate of China represents the value of Chinese currency – Yuan, it also partly sways the exchange rate and displays the power of China's economy on global economy. All in all, China's economic activity and real interest rate can become effective catalysts for forecasting future agricultural commodity price. According to the study by Klotz et al (2014), energy and industrial metals prices are Granger caused by China's economic activity at the 1% and 5% level of significance respectively. The relationship is primarily owing to China predominant position as the biggest importer of energy and metal sources such as oil, aluminum, and

copper in recent years (ibid). Though the paper excludes the causal impact of China's economy on other commodities' prices, the position of China in cotton market is relatively similar to that in energy and metal market. Therefore, it can be inferred of potential influence of China's economy on global cotton price.

Among all agricultural market, China has maintained the same position as the leading producer and consumer in cotton market for a long period of time. According to statista.com, China became the second top exporter and the top importer of cotton in 2016/2017 (n.d.). While level of cotton production of China is relatively equal to those of India and U.S., its cotton consumption level has increased overtime and had a significant impact on global cotton price. For instance, in the study of cotton price prediction in U.S. market, China's cotton import level is added in the model to forecast the supply and demand pattern of cotton more accurately (Cho, 2006; Isengildina-Massa and MacDonald, 2015). Moreover, between 2011 and 2013, as the cotton prices fell below the set level during the financial crisis in 2008, the Chinese government enforced a new policy to stockpile the domestic cotton in order to support the textile industry. The abruptly significant increase in cotton import of China raised the both its domestic and global cotton price substantially (Avins, 2015). Though the policy stopped in 2013, it still engendered a massive global cotton reserve with 60% in China in 2014, which in turn adversely affect the future price and future production of cotton. Therefore, it can be deduced that cotton import level of China has a negative influence over the global cotton price.

Overall, China's economic and monetary policies have great impact on global cotton price. Studies by Frankel (2008), and Klotz et al (2014) may suggest a negative correlation between China's real interest rate and global cotton price while according to research by Cho (2006), and Isengildina-Massa and MacDonald (2015), China's cotton import displays similar relation to global cotton price. Therefore, we testify the following hypotheses:

Hypothesis 3: Cotton price returns are negatively associated with China's imports

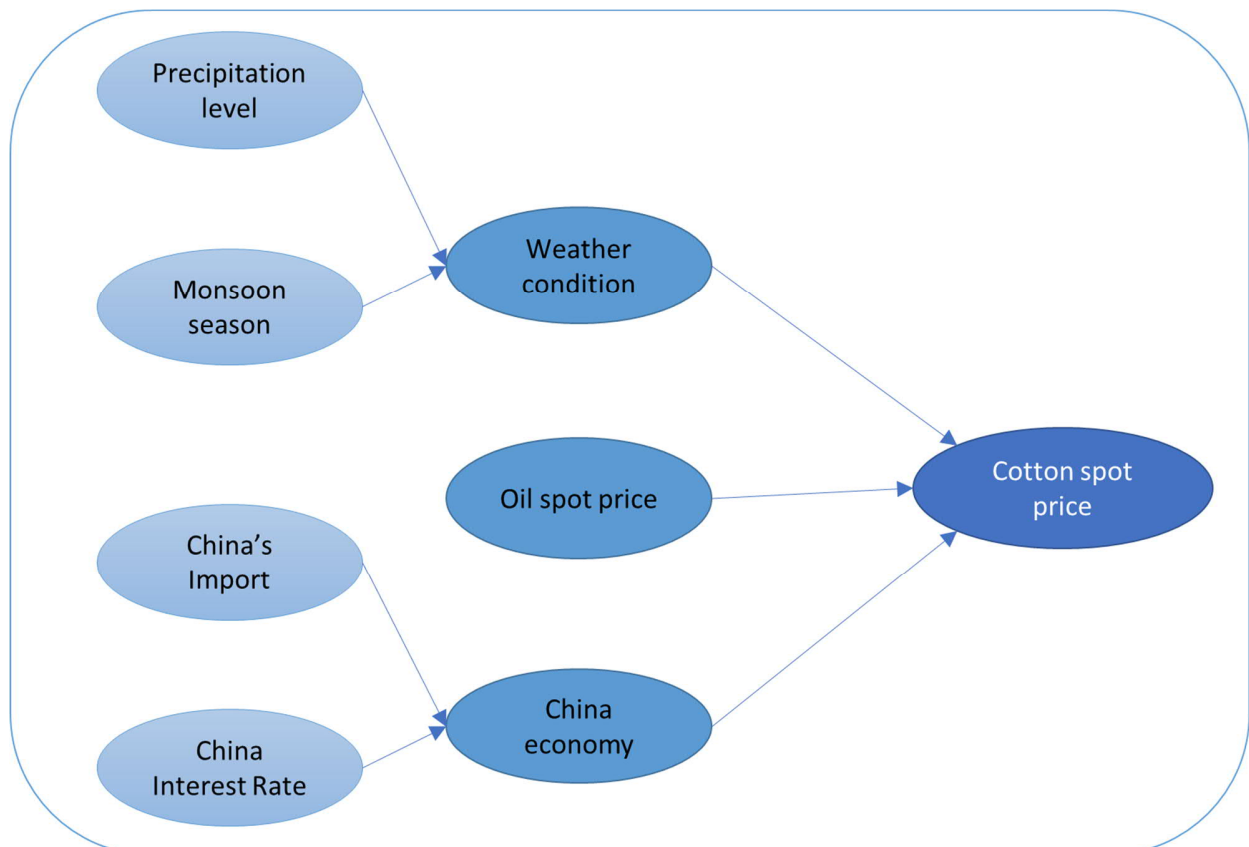
Hypothesis 4: Cotton price returns are negatively associated with China's interest rates

2.3 Summary

In conclusion, based on past studies, it can be deduced that weather factors, oil price and China's economy have a close relationship with cotton price. While oil price is believed to be proportional to the cotton price (Baffe, 2007), China's import level and China's real interest rate are inversely correlated to the cotton price (Klotz et al, 2014). In addition, weather conditions and monsoon seasons are expected to affect the cotton price, yet lacking thorough researches. Accordingly, additional tests should be conducted to examine the impact of weather condition, oil price and China's economy on global cotton price. Four hypotheses will be examined:

2.4 Conceptual Framework

Figure 1: Conceptual Framework



Based on previous findings and frameworks, four hypotheses are proposed:

Hypothesis 1: Cotton price returns are negative associated with monsoon (rainy) conditions

Hypothesis 2: Cotton price returns are positively associated with oil price return

Hypothesis 3: Cotton price returns are negatively associated with China's imports

Hypothesis 4: Cotton price returns are negatively associated with China's interest rate

3. METHODOLOGY

3.1 Theoretical model

The model is estimated based on the ordinary least squares method (OLS), which is widely employed to solve a spectrum of problems in econometrics. Ordinary least square (OLS) regression is a generalized linear modelling technique that can be employed to model a single response variable which has been recorded on at least an interval scale (Hayashi, 2000). In OLS method, the parameters of a linear function are selected with an aim to minimizing the sum of the squares of the errors (residuals) between the value of the observed dependent variable and the value of predicted dependent variable by the linear function (ibid). The application of technique is versatile, as it can be leveraged for single and multiple explanatory variables, and even categorical explanatory variables that have been pertinently coded. In the case of a model with n explanatory variables, the ordinary least squares regression model can be written as followed:

$$y = f(X) = f(x_1, x_2, x_3, x_4, \dots, x_n) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \dots + \beta_n x_n + \varepsilon$$

In the formula, y is denoted as the dependent variable, β_0 , is the intercept of the model, X_i corresponds to the i^{th} explanatory variable of the model ($i = \overline{1, n}$), and ε is the random error with expectation 0 and variance σ^2 .

As $f(X)$ is a linear function, it is deduced from the calculus differential equation:

$$\frac{\partial y}{\partial x_i} = \beta_i \text{ with } i = \overline{1, n}$$

Therefore, with other variables remaining constant, the coefficient β_i is the average expected change in value of y due to a change in value of x_i with i running from 1 to n .

3.2 Assumption of model

Though ordinary least squares (OLS) method is computationally feasible and can be applied in various econometrics test, it is vital to comprehend and examine the underlying assumptions of OLS regression. Should the assumption requirements be hardly fulfilled, the test will be misused and will provide invalid solutions. There are four principal assumptions for OLS multiple regression: normality of errors, correlation and multicollinearity, linearity and homoscedasticity (Hayashi, 2000; Osborne & Waters, 2002; Wooldridge, 2012).

Firstly, the observed data should have normal distribution of errors. The errors, also called residuals, is the difference between the observed value and the predicted value of dependent variable by the linear function. Non-normality of residuals can misinterpret the confidence intervals for model forecasting capability and create skewed distribution by increasing the emergence of outliers.

Secondly, the linear relationship should exist between the observed dependent and independent variables. The forecast value of dependent variable should be determined as a linear function of each independent variable, providing the other independent variables fixed. For each line, the slop should be unaltered by the values of the other variables. In addition, the influences of various independent variables on the expected value of the dependent variable are additive, or else the estimation will become correspondingly erroneous, especially when the model is used for the extrapolation beyond the sample data.

Thirdly, it is necessary to testify whether high correlation emerges among variables. A correlation matrix should be formed to check the existence of significant correlation, which subsequently generates multicollinearity. The multicollinearity can inflate the estimated variances, which makes it difficult to calculate the regression coefficients and standard

deviations accurately. In addition, inflated variances undermine the integrity of the regression analysis, interpretation, and conclusion. The common evaluation of multicollinearity is the variance inflation factor (VIF). When VIF is lower than 10, it implies no multicollinearity in the set of data.

The final requirement of OLS multiple regression is test of homoscedasticity of errors. Heteroscedasticity of errors can impair the reliability of standard errors of the OLS estimates and alter the confidence intervals to be either too narrow or too wide. Moreover, the violation of this assumption is inclined to place too much weight on some portion (subsection) of the data. Accordingly, error variances should be inspected to be constant across time for time series data and across independent variables.

3.3 Empirical model

The OLS model has been applied in previous commodity pricing studies such as the research by Espinosa and Goodwin on Kansas wheat pricing (1991), and Djunaidi's paper on Indonesian cocoa price estimation (1993). As for the cotton price estimation, the paper written by Ethridge and Davis (1982) and Cho (2006) also employed the model to estimate the cotton price based on its attributes and other factors such as weather, supply factors within the US market. Since this paper shares similarities with Cho's analysis (2006), the OLS multiple regression model is relatively pertinent to examine the relationship between cotton price with the supply and demand factors.

There are several complicated tests in OLS regression model. Djunaidi (1993) employed unit root test, Johansen's cointegration test and Granger's causality test for predicting Indonesian daily cocoa prices. Cho (2006) created composite cotton quality indices and conducted one-stage and two-stage hedonic models with R squares tests, F tests, Durbin-Watson test, and White test to predict the price of US cotton based on its characteristics, and supply factors. Nevertheless, this quantitative analysis is subjected to limitations owing to the narrow scope of a bachelor's thesis, as well as the constraint of budget and analytical tools. The assumption tests for multiple regression are presented to

examine criteria of linearity, normality of errors, independence and homoscedasticity, followed by the results of the multiple regression and its analysis.

3.4 Data and Variable selection

The study employs secondary data retrieved from Datastream (Thomson Reuters) and Climate Change Knowledge Portal (CCKP) database on a monthly basis from February of 1994 to December of 2015. The information of global cotton spot price, oil spot price, China's import, China's interest rate and monsoon period in India are collected from Datastream while data of precipitation level in China, India, and USA are extracted from CCKP.

Regarding dependent variable, historical cotton spot price data is retrieved from the S&P GSCI index (S&P Goldman Sachs Commodity Index) and is recorded in US dollar currency. In comparison, different independent variables are employed to demonstrate oil price, weather condition and China's economy. The historical average spot price of crude oil is recorded in US dollars to represent the overall impact of oil price. In terms of weather condition, the monthly precipitation level in USA, China and India are extracted in millimetres. Moreover, the monsoon data is also employed to better present the monthly impact of weather on cotton price. For each year, the monsoon value equals to 1 during monsoon period (from June to September) and equals to 0 in the remaining months. In addition, China's economy and currency value is represented by China's interest rate, and the relationship between China and world cotton market is demonstrated in China's import, recorded in percentage and US dollars respectively.

Further adjustments are conducted to incorporate the data into the model. The data are converted to difference between consecutive observations except for China's interest rate and monsoon. The primary justification is that the cotton price is hardly affected by the difference of China interest but the whole value of the contemporaneous interest rate. Moreover, the consecutive difference in monsoon value is meaningless and scarcely applicable into the analysis. Two types of consecutive difference are generated to analyse the data more thoroughly: natural logarithm and percentage difference. In natural

logarithm difference case, the value of difference at observation t equals to the changes between the original values between observation t and (t-1). In comparison, the percentage difference equals to the natural logarithm difference at observation t divided by the original value at observation (t-1). A table of types of variables is generated to explain the kind of variable in two case more easily.

Figure 2: Table of types of variables in both natural logarithm and percentage difference cases

		Variable's value in each case	
		Natural logarithm difference case	Percentage difference case
Dependent variable	Cotton spot price	Monthly logarithm change	Monthly percentage change
Independent variables	Oil spot price		
	China import		
	Rain level in China		
	Rain level in USA		
	Rain level in India		
	China interest rate	Monthly rate	Monthly rate
	Monsoon period	Binary (0 or 1)	Binary (0 or 1)

3.5 Suggested empirical equation

3.5.1 Natural logarithm difference

The regression function in this case can be written as followed:

$$D(\text{cotton})_t = \beta_0 + \beta_1 * D(\text{oil})_t + \beta_2 * D(\text{China Import})_t + \beta_3 * \text{China Interest Rate}_t \\ + \beta_4 * D(\text{Rain in USA})_t + \beta_5 * D(\text{Rain in China})_t + \beta_6 * D(\text{Rain in India})_t + \beta_7 \\ * \text{Monsoon} + \varepsilon_t$$

Where

- $D(\text{cotton})_t$ is the natural logarithm change of cotton price between observation t and (t-1)
- β_0 is the intercept of the function
- $\beta_1, \beta_2, \beta_3, \beta_4, \beta_5, \beta_6, \beta_7$ is the beta coefficient of the corresponding independent variables
- $D(\text{oil})_t$ is the natural logarithm change of oil price between observation t and (t-1)
- $D(\text{China Import})_t$ is the natural logarithm change of China's Import level between observation t and (t-1)
- $\text{China Interest Rate}_t$ is the China's interest rate at observation t
- $D(\text{Rain in USA})_t$ is the natural logarithm change of USA's rainfall level between observation t and (t-1)
- $D(\text{Rain in China})_t$ is the natural logarithm change of China's rainfall level between observation t and (t-1)
- $D(\text{Rain in India})_t$ is the natural logarithm change of India's rainfall level between observation t and (t-1)
- Monsoon equals to 1 if monsoon season occurs or else equals to 0 at observation t
- ε_t is the error term at observation t

3.5.2 Percentage difference

The regression function in this case can be written as followed:

$$\%D(\text{cotton})_t = \beta_0 + \beta_1 * \%D(\text{oil})_t + \beta_2 * \%D(\text{China Import})_t + \beta_3 * \text{China Interest Rate}_t \\ + \beta_4 * \%D(\text{Rain in USA})_t + \beta_5 * \%D(\text{Rain in China})_t + \beta_6 * \%D(\text{Rain in India})_t \\ + \beta_7 * \text{Monsoon} + \varepsilon_t$$

Where

- $\%D(\text{cotton})_t$ is the percentage change of cotton price at observation t over that at observation (t-1)
- β_0 is the intercept of the function

- $\beta_1, \beta_2, \beta_3, \beta_4, \beta_5, \beta_6, \beta_7$ is the beta coefficient of the corresponding independent variables
- $\%D(\text{oil})_t$ is the percentage change of oil price at observation t over that at observation (t-1)
- $\%D(\text{China Import})_t$ is the percentage change of China's Import level at observation t over that at observation (t-1)
- China Interest Rate_t is the China's interest rate at observation t
- $\%D(\text{Rain in USA})_t$ is the percentage change of USA's rainfall level at observation t over that at observation (t-1)
- $\%D(\text{Rain in China})_t$ is the percentage change of China's rainfall level at observation t over that at observation (t-1)
- $\%D(\text{Rain in India})_t$ is the percentage change of India's rainfall level at observation t over that at observation (t-1)
- Monsoon equals to 1 if monsoon season occurs or else equals to 0 at observation t
- ε_t is the error term at observation t

4. FINDINGS

4.1 Descriptive statistics

4.1.1 Natural log difference

Figure 3: Descriptive statistics table in the case of natural logarithm difference

	D(cotton)	D(oil)	D(China Import)	China Interest Rate	D(Rain in USA)	D(Rain in China)	D(Rain in India)	Monsoon
Number of observations	263	263	263	263	263	263	263	263
Number of missing	0	0	0	0	0	0	0	0
Mean	-0.0002	0.0044	12.1386	0.0682	0.0018	0.0033	-0.0045	0.3346
Standard Error	0.0055	0.0058	0.3155	0.0012	0.0138	0.0345	0.0564	0.0291

Median	0.0020	0.0069	13.6716	0.0600	0.0091	0.0830	0.0075	0
Minimum	-0.3485	-0.3320	0	0.0435	-0.8897	-1.5355	-2.1967	0
Maximum	0.2766	0.8626	17.6121	0.1206	0.6349	1.5144	3.4460	1
Range	0.6252	1.1947	17.6121	0.0771	1.5246	3.0500	5.6428	1
Standard Deviation	0.0894	0.0943	5.1175	0.0202	0.2243	0.5607	0.9152	0.4727
Sample Variance	0.0080	0.0089	26.1895	0.0004	0.0503	0.3143	0.8377	0.2235
Kurtosis	1.7861	26.2174	1.5506	0.6734	1.3248	-0.2160	0.0663	-1.5144
Skewness	-0.2475	2.2794	-1.6925	1.4307	-0.2207	-0.2641	0.0622	0.7050

It is inferred from the table that D(cotton), D(rain in China) and D(rain in India) have relatively symmetric distribution with skewness fitting the normal distribution range between $-1/2$ and $1/2$ (Bulmer, 1979). In contrast, other independent variables are highly-skewed distributed with absolute value of skewness larger than 1 except for the monsoon value with skewness of 0.705 (ibid). In terms of kurtosis, the kurtosis of most variables is lower than 3, indicating shorter and thinner tails an overall lower and broader central peak compared to normal distribution (Westfall, 2014). The only exception is that of D(oil), presenting the opposite pattern to that of the former.

4.1.2 Percentage difference

Figure 4: Descriptive statistics table in the case of percentage difference

	%D(cotton)	%D(oil)	%D(China Import)	China Interest Rate	%D(Rain in USA)	%D(Rain in China)	%D(Rain in India)	Monsoon
Number of observations	263	263	263	263	263	263	263	263
Number of missing	0	0	0	0	0	0	0	0

Mean	0.0037	0.0093	12.1386	0.0682	0.0270	0.1641	0.5419	0.3346
Standard Error	0.0055	0.0070	0.3155	0.0012	0.0142	0.0401	0.1412	0.0291
Median	0.0020	0.0069	13.6716	0.0600	0.0091	0.0866	0.0076	0
Minimum	-0.2943	-0.2825	0	0.0435	-0.5892	-0.7846	-0.8888	0
Maximum	0.3187	1.3694	17.6121	0.1206	0.8869	3.5468	30.3766	1
Range	0.6130	1.6520	17.6121	0.0771	1.4761	4.3315	31.2654	1
Standard Deviation	0.0893	0.1132	5.1175	0.0202	0.2311	0.6505	2.2905	0.4727
Sample Variance	0.0080	0.0128	26.1895	0.0004	0.0534	0.4231	5.2466	0.2234
Kurtosis	1.5197	79.2301	1.5506	0.6734	1.5216	5.4166	111.5581	-1.5144
Skewness	0.2286	6.4247	-1.6925	1.4307	0.7709	1.6343	9.0450	0.7051

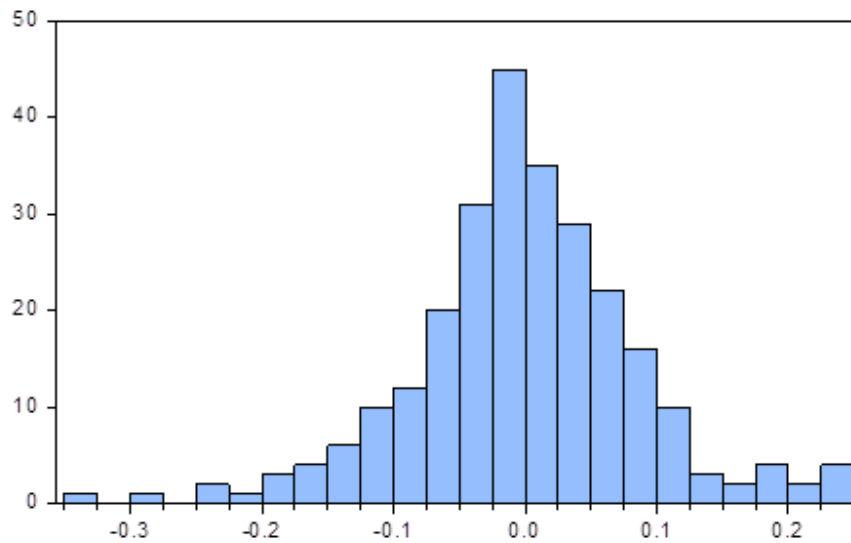
It is inferred from the table that only the dependent variable - D(cotton) - is comparatively symmetrically distributed, followed by %D(rain in USA) and the monsoon with moderately skewed distribution. However, other independent variables have skewness with absolute value larger than 1, suggesting highly skewed distribution of data. Moreover, while the kurtosis of most variables is lower than 3, that of %D(oil) and %D(rain in India) are substantially higher than 3, signaling longer and fatter tails, higher and sharper central peak in comparison to a normal distribution.

4.2 Normality of error test

To test the assumption of normality of error, a histogram of the residual value of the regression is formed. The residual is calculated as the error (difference) between observed values (observed cotton spot prices) and predicted values in the same period.

4.2.1 Natural log difference

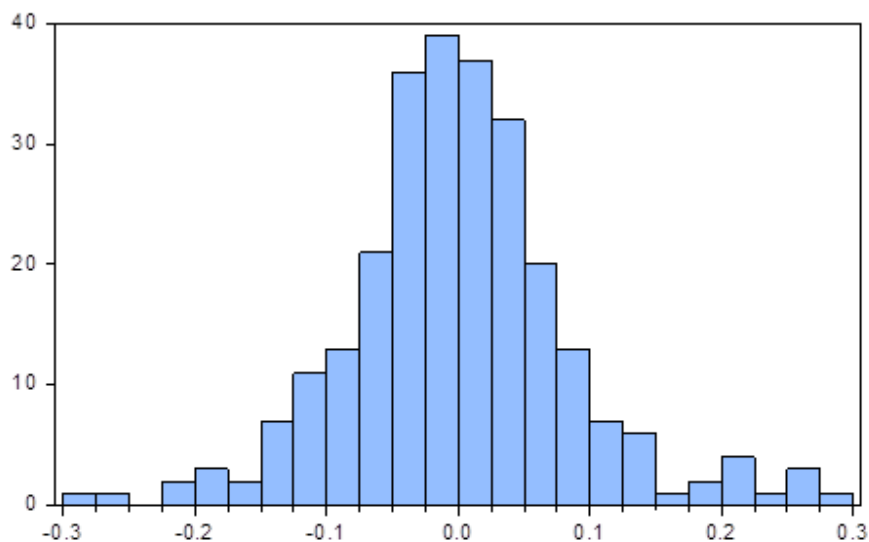
Figure 5: Histogram of residuals of regression in the case of natural logarithm difference



The histogram has a fitted normal curve with skewness between $-1/2$ and $1/2$, satisfying the criteria of normality of errors.

4.2.2 Percentage difference

Figure 6: Histogram of residuals of regression in the case of percentage difference

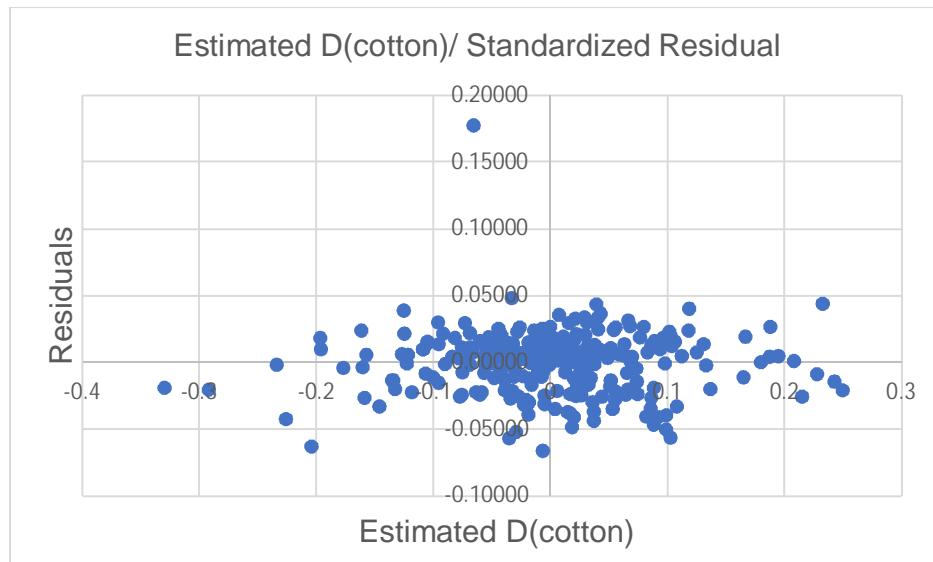


In this case, the histogram shows the similar pattern to that in former case, with skewness of 0.29 within the normal distribution skewness range. Therefore, the assumption of normal distribution of errors is fulfilled.

4.3 Linearity test

4.3.1 Natural log difference

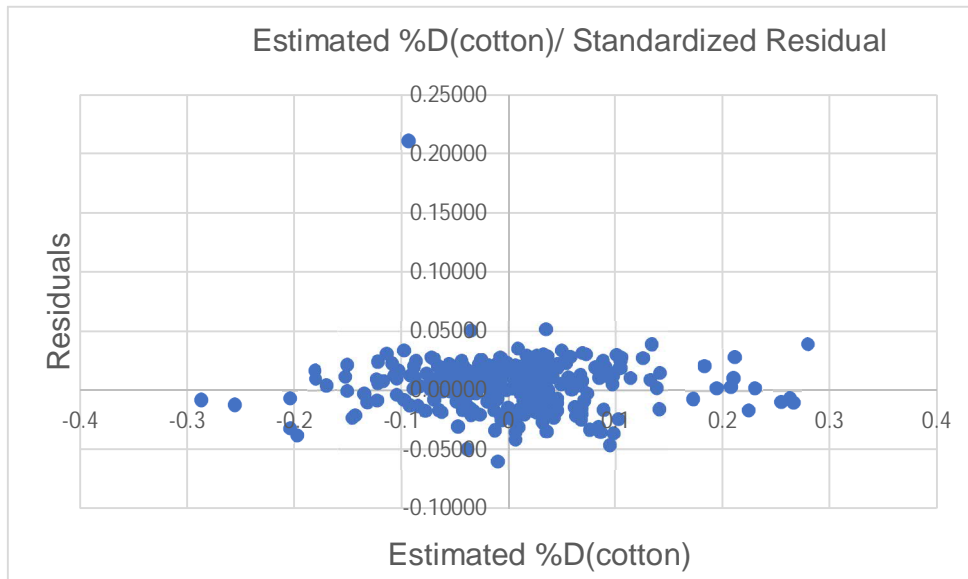
Figure 7: Scatter plot of predicted D(cotton) and standardized residual



The residuals scatter roughly symmetrically around the predicted natural logarithm change in cotton spot price line, which could imply the linearity of the regression model.

4.3.2 Percentage difference

Figure 8: Scatter plot of predicted %D(cotton) and standardized residual



The scatter plot in the case of percentage difference displays similar patterns with that in the instance of natural logarithm difference. As the result, the criteria of linearity can be gratified for further analysis for both cases.

4.4 Correlation test – multicollinearity

4.4.1 Natural log difference

Figure 9: Correlation matrix among all variables in the case of natural logarithm change

	D(cotton)	D(oil)	D(China Import)	China Interest Rate	D(Rain in USA)	D(Rain in China)	D(Rain in India)	Monsoon
D(cotton)	1							
D(oil)	0.2045	1						
D(China Import)	-0.0552	0.0152	1					
China Interest Rate	-0.0223	-0.0059	-0.7876	1				

D(Rain in USA)	-0.0031	0.0186	-0.0353	-	1			
			0.0023					
D(Rain in China)	-0.0463	0.0971	-0.0691	0.0277	0.2784	1		
D(Rain in India)	-0.0245	0.1335	-0.0178	0.0161	0.1453	0.5599	1	
Monsoon	-0.0968	0.0421	0.0743	0.0008	-0.0134	-0.0344	0.2278	1

The majority of correlation values line between -0.09 and 0.09, which implies that those independent variables are not correlated to each other. However, the change in China's import and value of China's interest rate is significantly negatively correlated (-0.7876) due to the impact of interest rate on exchange rate. Moreover, while the precipitation level in USA is moderately correlated to that in China and India, the correlational value between the latter countries is 0.5599, which can be justified by the similar weather condition in the same continent.

In correlation with natural logarithm change in cotton spot price, the change in oil price accounts for the highest absolute correlation value of 0.2045 whereas other independent variables hold lower value ranging between -0.1 to 0.1.

Figure 10: Multicollinearity statistics table in the case of natural logarithm change

Variable	Coefficient Variance	Uncentered VIF	Centered VIF
Constant	0.0013	66.9899	NA
China Interest Rate	0.0978	28.013	2.4533
Monsoon	0.0001	1.6737	1.2131
D(China Import)	1.91E-06	15.1349	2.3954
D(oil)	0.0024	1.2540	1.2340
D(Rain in China)	0.0001	2.2346	2.2332

D(Rain in India)	3.11E-05	1.9116	1.8691
D(Rain in USA)	0.0003	1.1822	1.1821

In order to better examine the multicollinearity of the data, the Variance Inflation Factor (VIF) ratio is calculated as shown in Figure 10. The uncentered VIF excludes the constant variable from the equation, therefore creates less reliable evidence in this model. The VIFs of all variables are lower than 10, implying no multicollinearity in the case of natural logarithm difference.

4.4.2 Percentage difference

Figure 11: Correlation matrix among all variables in the case of percentage change

	%D(cotton)	%D(oil)	%D(China Import)	China Interest Rate	%D(Rain in USA)	%D(Rain in China)	%D(Rain in India)	Monsoon
%D(cotton)	1							
%D(oil)	0.1861	1						
%D(China Import)	-0.0567	0.0351	1					
China Interest Rate	-0.0243	-0.0198	-0.7876	1				
%D(Rain in USA)	0.0192	0.0117	-0.0317	-0.0135	1			
%D(Rain in China)	-0.0503	0.0464	-0.0848	0.0503	0.2570	1		
%D(Rain in India)	-0.0281	0.0065	0.0090	-0.0135	0.0553	0.2553	1	
Monsoon	-0.0884	0.0193	0.0744	0.0008	-0.0607	-0.1228	0.0335	1

The correlation table shows the familiar pattern to that in the case of natural logarithm difference. Though most values are between the low correlation range, the figures of precipitation among three countries are comparatively moderate between 0.0553 to 0.2570. Furthermore, the percentage change in China's import is reversely linked to the value of China's interest rate, showing similar results with that of natural logarithm change case.

In correlation with the percentage change in cotton spot price, the pattern follows that in the former case, yet the absolute correlation value in the case of percentage change is lower.

Figure 12: Multicollinearity statistics table in the case of percentage change

Variable	Coefficient Variance	Uncentered VIF	Centered VIF
C	0.0023	80.0927	NA
%D(oil)	0.0023	1.0112	1.0044
%D(China Import)	3.01E-06	17.8963	2.6921
China Interest Rate	0.1905	33.1012	2.6679
Monsoon	0.0001	1.5544	1.0343
&D(Rain in China)	8.09E-05	1.2445	1.1698
&D(Rain in India)	6.00E-06	1.1361	1.0757
&D(Rain in USA)	0.0006	1.0898	1.0752

It is demonstrated in Figure 12 that all of the centred VIFs are lower than 10, implying no multicollinearity in the regression model in the case of percentage change. Overall, both cases meet the requirement of multicollinearity of the model.

4.5 Homoscedasticity test

4.5.1 Natural log difference

Figure 13: Standardized residual versus predicted D(cotton) over time

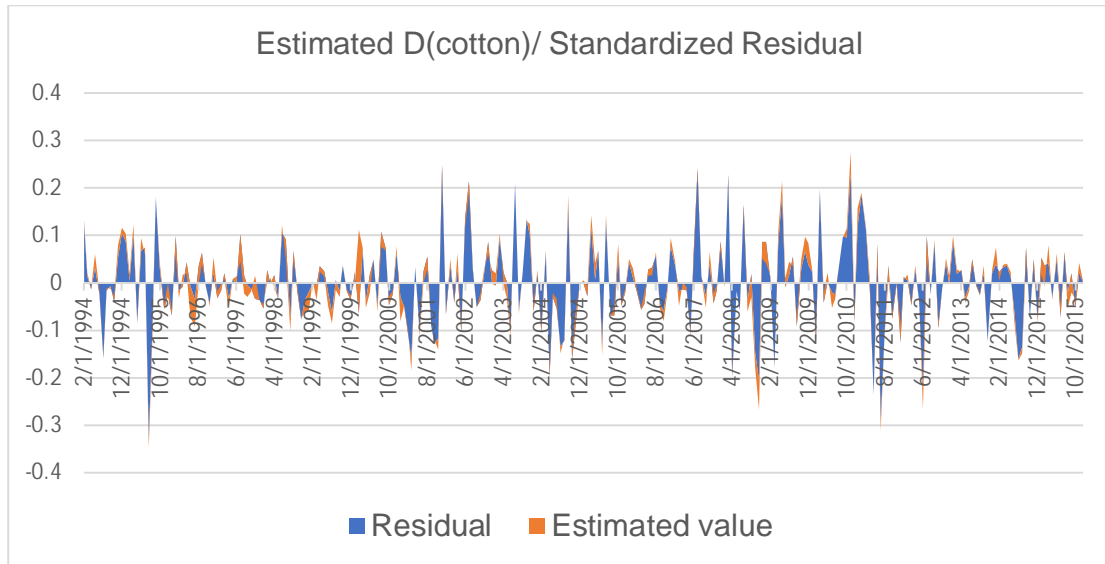
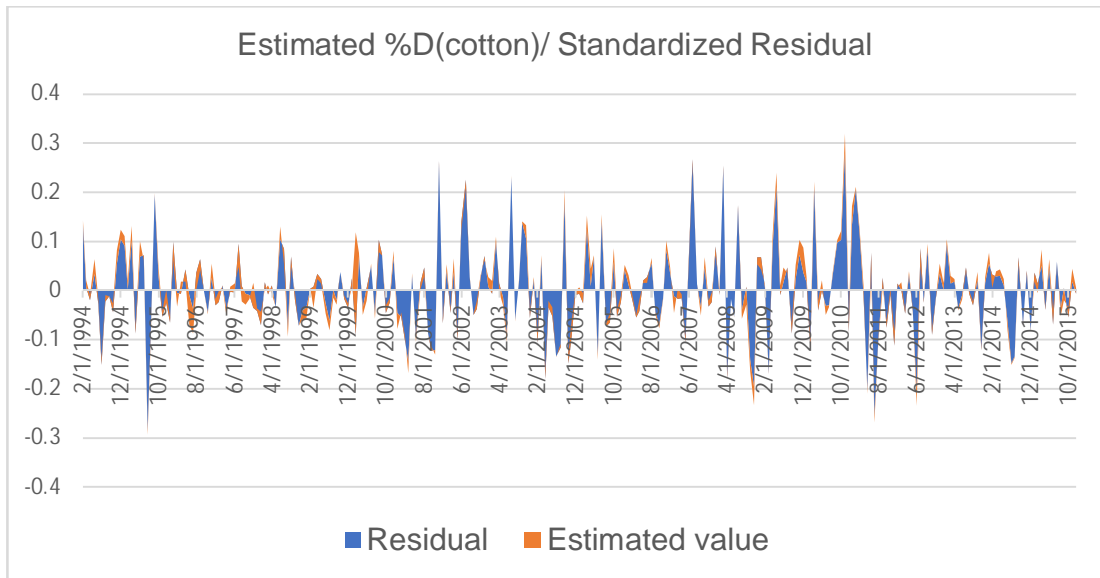


Figure 7 and Figure 13 indicate the consistency of the variance of the residuals and the independence of the residuals from independent variables, therefore the data is homoscedastic in the case of natural logarithm difference.

4.5.2 Percentage difference

Figure 14: Standardized residual versus predicted %D(cotton) over time



It is inferred from Figure 8 and Figure 14 that the variance of the residuals remains constant and the standardized residuals of percentage change in cotton price is unaffected by independent variables. Correspondingly, both cases attain the homoscedasticity requirement for further analysis of the multiple regression model.

5. DISCUSSION AND ANALYSIS

5.1 Natural log difference

Figure 15: Empirical result table in the case of natural logarithm difference

Variable	Coefficient	Std. Error	t-Statistic	Prob.
Constant	0.0957	0.0372	2.5691	0.0108
China Interest Rate	-0.7396	0.3127	-2.3651	0.0188
Monsoon	-0.0187	0.0119	-1.5744	0.1166
D(China Import)	-0.0033	0.0014	-2.3864	0.0177
D(oil)	0.2055	0.0488	4.2080	0.0001
D(Rain in China)	-0.0146	0.0122	-1.1977	0.2321
D(Rain in India)	0.0018	0.0055	0.3281	0.7431
D(Rain in USA)	0.0029	0.0172	0.1694	0.8656

R-squared	0.0718	Mean dependent variable	-0.0002
Adjusted R-squared	0.0463	S.D. dependent variable	0.0894
S.E. of regression	0.0873	Log likelihood	272.0317
Sum squared residual	1.9456	Durbin-Watson stat	2.1620
Prob(F-statistic)	0.0075	F-statistic	2.8207
Prob(Wald F-statistic)	0.0011	Wald F-statistic	3.5851

The result of regression in the case of natural logarithm difference is shown in Figure 15. The probability of F statistics is lower than alpha level ($0.0075 < 0.05$), signalling the significance of the result. The R-squared and adjusted R-squared value are relatively small (0.0718 and 0.0463 respectively), indicating that only 4.63% of D(cotton)'s movement is explained by the movement of independent variables. The Durbin-Watson statistics is comparatively close to 2 ($2.1620 > 2$), suggesting a weakly negative autocorrelation in the residuals in the regression analysis.

Among all coefficient, China's interest rate accounts for the highest coefficient in absolute value, and it becomes the most influential to natural logarithm change in cotton global spot price. With the p-values lower than the alpha value of 0.05, China's interest rate and the change in China's import level are proven to have a strong regression with change in the cotton price. The result can be justified by the position of China in global cotton market. China remains the largest cotton importing country and the top three producing countries for many years, generating a similar event to its dominant power the copper market (Klotz et al, 2014). Since China's interest rate can directly affect the exchange rate of Yuan to other currencies, and therefore substantially influences the global price of cotton. In addition, the cotton stock-piling policy in China has accumulated a considerable margin within the world total import levels so that China's import level can manipulate the demand and supply of cotton to a degree. Indeed, the negative coefficient of China interest rate and change in China's import indicates the reverse relationship between the former two variables and change in global cotton price. The results reassure the arguments of relationship between China's economy and global cotton price in the literature review, signaling a hazardous monopoly in the market. Hedgers, usually farmers

and producers, should take into consideration the recent cotton news and policies in China in order to better predict future price movement. Furthermore, international organization, and policy makers, particularly from top cotton exporting countries such as India or USA, should propose new enforcements to reduce the power of China over the market.

Another point worth consideration is the change in oil price. Since its P-value is lower than 0.05 and coefficient equals to 0.2055, the change in the oil price is demonstrated to earn a notably positive link to the change in cotton price. The association can be primarily explained by the high percentage of oil cost in production cost of cotton. Moreover, during the financial crisis period from 2007 to 2009, the volatility in oil price posed a spillover effect on other markets, including cotton market (Ji and Yang, 2012). After the crisis, the effect lingered and strengthened the association between oil and cotton. The data supported Baffe's analysis (2005) and offers suggestion that hedgers should partly observe the movement of oil price to estimate the cotton price more precisely.

Precipitation variables exhibit different results in the regression test. The change in rainfall in China generates a negative coefficient, which is opposite to those in India and USA. Correspondingly, it is deduced that the global price of cotton will accelerate when the rainfall level in China diminishes and that in India and USA increases. The data is contradictory against the intuitive assumption of similarity among three weather variables. In addition, the p-value of all three variables are significantly higher than the alpha value of 0.05, heralding the weak evidence of impact on cotton natural logarithm change in price. The weak significance of coefficients might be associated with the low frequency of employed data as monthly one, whereas weekly or daily rainfall usually create temporary impact on the growing process of cotton. Since monthly data can hardly incorporate weather information quite simultaneously, the influence of monthly rainfall level over change in cotton price is alleviated. Additionally, the analyzed precipitation information is extracted for the whole countries rather than only the cotton producing regions, consequently diminishing the power over cotton price. In contrast to the weak forecasting power of three variables, the monsoon variable has a better P-value of 0.1166 and coefficient of -0.018, implying its negative impact on changes in cotton price at 85%

confidence. The discrepancy might be attributed to the frequency of monsoon data, since monsoon happens monthly, which corresponds to the monthly data of cotton price.

5.2 Percentage difference

Figure 16: Empirical result table in the case of percentage difference

Variable	Coefficient	Std. Error	t-Statistic	Prob.
Constant	0.1041	0.0483	2.1551	0.0321
%D(oil)	0.1537	0.0479	3.2098	0.0015
%D(China Import)	-0.0035	0.0017	-2.0138	0.0451
China Interest Rate	-0.7661	0.4364	-1.7554	0.0804
Monsoon	-0.0161	0.0116	-1.3885	0.1662
%D(Rain in China)	-0.0113	0.0090	-1.2594	0.2090
%D(Rain in India)	-0.0003	0.0024	-0.1174	0.9066
%D(Rain in USA)	0.0095	0.0243	0.3920	0.6954
R-squared	0.0642	Mean dependent variance		0.0037
Adjusted R-squared	0.0385	S.D. dependent variance		0.0893
S.E. of regression	0.0876	Durbin-Watson stat		2.1443
Sum squared residual	1.9555	F-statistic		2.5009
Log likelihood	271.3624	Prob(F-statistic)		0.0167

The result of regression in the case of percentage difference is displayed in Figure 16. The probability of F statistics is lower than alpha level ($0.0167 < 0.05$), heralding the significance of the result. The R-squared and adjusted R-squared value are relatively small (0.0642 and 0.0385 respectively), indicating that only 3.85% of %D(cotton)'s movement is explained by the movement of independent variables, which is lower than that in the case of natural logarithm difference. With the Durbin-Watson statistics slightly higher than 2 ($2.1443 > 2$), the analysis has a weakly negative autocorrelation in the residuals, presenting a similar result in the case of natural logarithm difference.

Among all independent variables, only percentage change in China's import and percentage change in oil price have the P-value lower than alpha, implying strong impact on percentage change in cotton price. The result is parallel to that in the case of natural logarithm difference, reinforcing their correlation with the global cotton price. In comparison, the P-value of China's interest rate is 0.0804, reducing its impact to only at 90% confidence.

As regards the precipitation variables, only the coefficient in percentage rainfall change in India changes to be negative while that in China and USA remain the same sign. In other words, the global price of cotton will display an upward trend as the rainfall level in China and India decreases and that in USA increases. Moreover, all three coefficients are of weak significance, which can be associated to similar justification in the case of natural logarithm difference. In addition, in this case, the p-value of monsoon increases, signaling weaker evidence of impact on percentage change in cotton price.

6. CONCLUSIONS

6.1 Main Findings

The study examines the impact of weather factors, oil price return and China's economy on the cotton price return. Prior analyses claimed the negative relationship between cotton price and oil price, China's import and China's interest rate. As for weather variables, there has been hardly conclusions between monsoon period and the cotton price return. In order to testify the hypotheses, a multiple regression model is conducted for 263 samples of cotton spot price and independent variables: oil spot price, China's import, China's interest rate, precipitation in USA, China, India and monsoon period, which are all recorded as monthly data from February of 1994 to December of 2015.

The assumption for multiple regression model is examined. Normal distribution histogram for residuals affirms the normality assumption. Pearson's correlation matrix and multicollinearity test could satisfy the no-collinearity assumption. The scatterplot of predicted cotton price return versus residuals rationalizes the linearity of the model. The

plot of residuals over time supports the homoscedasticity of the data set.

Correspondingly, the input set of data can be considered valid for conducting multiple regression model.

The multiple regression model reveals that approximately 4% of the movement cotton price return can be explained by the movements of the independent variables in the sample. The empirical results confirm the three out of four assumed hypotheses. Firstly, oil price return is positively correlated with cotton price return at 95% confidence. Moreover, both the change in China's import level and China's interest rate establish converse relationships with cotton price return at 95% confidence. As regards weather factors, the evidences are insufficient to prove link between rainfall change in China, India, USA and cotton price return. In addition, monsoon period generates a negative link with the natural logarithm cotton price return at 85% confidence yet failed to prove the similar pattern with percentage cotton price return. Among all independent variables, China's interest rate has the highest coefficient in absolute value in both cases, asserting the strong negative link between China's interest rate and cotton price return.

This paper reaffirms the potential influences on the cotton spot price besides cotton futures price: China's interest rate, China's import and oil spot price. Hedgers can take into consideration the mentioned variable with an aim to forecasting future cotton price movement more precisely. Farmers and manufacturers can adjust their operations to the market changes with more confidence.

6.2 Limitation of the research

Despite the remarkable result of the regression analysis, the significant correlation between cotton price return and other variables, comprising oil price return, China's interest rate, and changes in China's import level, fails to prove causation of cotton price return by those variables. Correlation merely testifies the variables at a linear relationship, which may overlook the underlying cause-and-effect relationship of the data. Additionally, the short length of time series data (February 1994 to December 2015) and the incapacity

to acquire most recent data might adversely impact the accuracy and long-term application of the model. The scale of the thesis is restricted with only 263 samples and low frequency. Several independent variables are only recorded as monthly data, so other variables need converting into the same frequency.

The multiple regression model generates results based on the past value of independent variables, therefore the result can be inappropriately applied in the matter of future shocks. Moreover, the model only utilizes the China's import level as the demand variable and disregards other importing countries such as Vietnam, Pakistan, Bangladesh and United States.

6.3 Suggestions for Further Research

The multiple regression pricing model for cotton price return can be later adjusted for other commodities such as sugar, wheat and corn. They are major exporting and importing products for various countries and are quickly responsive to changes in market. Daily data are suggested for the regression model since they capture the influence of independent variables more effectively. Moreover, hedgers prefer short-term forecast to earn profits in short period of time. More advanced analysis tests and models such as co-integration test, causality test, and error correction model are needed to improve the performance and the integrity of the result.

REFERENCES

Akram, F. (2009) 'Commodity prices, interest rates and the dollar'. *Energy Economics*. [Online]. 31: 838–851.

Ashby, F. (1929) 'Cotton Futures as Forecasters of Cotton Spot Prices'. *Journal of the American Statistical Association* [Online]. 24(168): 412-419.

Avins, J (2015) 'Why has the world stockpiled enough cotton for 127 billion t-shirts?' Available from: <https://qz.com/573860/why-has-the-world-stockpiled-enough-cotton-for-127-billion-t-shirts/> [Accessed on 15th March 2018]

Aziz, M. and Bakar, N. (2013) 'Oil price fluctuation and changing comparative advantage'. *Sosyoekonomi*: 108-130.

Baffes, J. (2007) 'Oil Spills on Other Commodities'. Policy Research Working Paper 4333, World Bank.

Bange, M., Constable, G., McRae, D., Roth, G. (2010) 'Cotton' In: Stokes, C., Howden, M. (ed.) *Adapting Agriculture to Climate Change: Preparing Australian Agriculture, Forestry and Fisheries for the Future*. CSIRO. pp. 49–66.

Bhanumurthy, N., Pami Dua, P., and Kumawat, L. (2012) 'Weather Shocks, Spot and Futures Agricultural Commodity Prices: An Analysis for India'. Available from: <http://www.cdedse.org/pdf/work219.pdf> [Accessed on 15th March 2018]

Belles, J. (2017) 'Southwest Monsoon 2017 Forecast: Warmer-Than-Average Conditions Could Lead to More Storms'. Available from: <https://weather.com/news/weather/news/southwest-monsoon-2017-begin> [Accessed on 15th March 2018]

Bernanke., S. (2008) 'Outstanding Issues in the Analysis of Inflation'. In: *Federal Reserve Bank of Boston's 53rd Annual Economic Conference*, Chatham, Massachusetts: 9 June. Available from: <https://www.federalreserve.gov/newsevents/speech/bernanke20080609a.htm> [Accessed on 15th March 2018]

Blanc, E., & Quirion, P., & Strobl, E. (2008.) 'The climatic determinants of cotton yields: evidence from a plot in west Africa'. *Agric. For. Meteorol.* [Online]. 148 (6–7): 1093–1100.

Brooks, C., and Prokopczuk, M., and Wu, Y. (2013) 'Commodity futures prices: More evidence on forecast power, risk premia and the theory of storage'. *The Quarterly Review of Economics and Finance* [Online]. 53 (1): 73-85.

Brorsen, B., Bailey, D., and Richardson, J. (1984) 'Investigation of Price Discovery and Efficiency for Cash and Futures Cotton Prices'. *Western Journal of Agricultural Economics* [Online]. 9(1): 170-176.

Bulmer, G. (1979) *Principles of Statistics*. Dover.

Cashin P., Cespedes L., & Sahay R. (2003) 'Commodity currencies and the real exchange rate'. *Journal of Development Economics* [Online]. 75: 239-268.

Chang, H. & Su, M. (2010) 'The substitutive effect of biofuels on fossil fuels in the lower and higher crude oil price periods'. *Energy* [Online]. 35: 2807–2813.

Chapagain, A., & Hoekstra, A., & Savenije, H., & Gautam, R. (2006) 'The water footprint of cotton consumption: An assessment of the impact of worldwide consumption of cotton products on the water resources in the cotton producing countries'. *Ecological Economics* [Online]. 60(1):186-203.

Chavez, M. (2007) *The efficiency of the U.S. cotton futures market (1986-2006): normal backwardation, co-integration, and asset-pricing*. Master Thesis. Texas, U.S: Graduate Studies of Texas A&M University [published].

Chen Y.-C., Rogoff K., & Rossi B. (2008) 'Can exchange rates forecast commodity prices?' In: *NBER Working Paper No. 13901*. Cambridge, MA: National Bureau of Economic Research.

Chen, L., and Jackson, D., and Kim, H., and Resiandini, P. (2014) 'What drives commodity prices?'. *American Journal of Agricultural Economics* [Online]. 96(5), 1455–1468.

Cheung, W., and Chinn, D., and Pascual, G. (2005) 'Empirical exchange rate models of the nineties: are any fit to survive?'. *J. Int. Money Financials*. [Online]. 24: 1150–1175.

Cho, H. (2006) *An Innovative Hedonic Pricing System for US Cotton Based on HVI Quality Measures and Market Factors*. Doctor Dissertation. Raleigh, NC: Graduate Faculty of North Carolina State University. [published]

Constable, G. & Bange, M. (2015) 'The yield potential of cotton (*Gossypium hirsutum* L.)'. *Field Crops Research* [Online]. 182: 98-106.

Cotton production worldwide by top countries (n.d.). Available from: <https://www.statista.com/statistics/263055/cotton-production-worldwide-by-top-countries/> [Accessed on 15th March 2018]

Engel, C., and West, D. (2005) 'Exchange rates and fundamentals'. *Journal of Political Economics* [Online]. 113: 485–517.

Ethridge, D. & Davis, B. (1982) 'Hedonic Price Estimation for Commodities: an Application to Cotton'. Available from: <http://ageconsearch.umn.edu/record/32258/files/07020293.pdf> [Accessed on 15th March 2018]

Fadiga, M. and Misra, S. (2007) 'Common Trends, Common Cycles, and Price Relationships in the International Fiber Market'. *Journal of Agricultural and Resource Economics* [Online]. 23(1): 154-168.

Food and Agriculture Organization of the United Nations (2013) 'FAO Statistical Yearbook 2013- World Food and Agriculture'. Available from: <http://www.fao.org> [Accessed on 21st Feb 2018]

Frankel, J. and Rose, A. (1995) 'Chapter 33 Empirical research on nominal exchange rates'. *Handbook of International Economics, Elsevier* [Online]. 3: 1689-1729.

Gargano, A. and Timmermann, A. (2014) 'Forecasting commodity price indexes using macroeconomic and financial predictors'. *International Journal of Forecasting* [Online]. 30: 825–843.

Hanson, K., and Robinson, S., and Schluter, G. (1993) 'Sectoral effects of a world oil price shock: economywide linkages to the agricultural sector'. *Journal of Agricultural and Resource Economics* [Online]. 18 (1), 96–116.

Hayashi, F. (2000) *Econometrics*. Princeton University Press.

Hudson, D., Elam, E., Ethridge, D. & Brown, J. (1996) 'Price Information in Producer Markets: An Evaluation of Futures and Spot Cotton Price Relationships in the Southwest Region Using Cointegration'. *Agribusiness* [Online]. 12(4): 363.

Isard, P. (1995) *Exchange Rate Economics*. Cambridge: Cambridge University Press.

Ji, Q. and Fan, Y. (2012) 'How does oil price volatility affect non-energy commodity markets?'. *Applied Energy* [Online]. 89: 273–280.

Isengildina-Massa, O. & MacDonald, S. (2015) 'Forecasting U.S. Cotton Prices in a Changing Market'. *Advances in Business and Management Forecasting* [Online]. 9: 91-113.

Joëts, M. and Mignon, V. and Razafindrabe, T. (2017) 'Does the volatility of commodity prices reflect macroeconomic uncertainty?'. *Energy Economics* [Online]. 68: 313–326.

Kilian, L., and Taylor, P. (2003) 'Why is it so difficult to beat the random walk forecast of exchange rates?'. *Journal of International Economics* [Online]. 60 (1): 85–107.

Krishnamurti, T. (2015) 'Indian Monsoon'. Available from:

<https://www.britannica.com/science/Indian-monsoon> [Accessed on 15th March 2018]

Klotz, P. and Lin, T., and Hsu, S. (2014) 'Global commodity prices, economic activity and monetary policy: The relevance of China'. *Resources Policy*, 42: 1–9.

Leading cotton importing countries in 2016/2017 (n.d.). Available from:
<https://www.statista.com/statistics/191896/leading-cotton-importing-countries/> [Accessed on 15th March 2018]

Mark, N. (2001) *International Macroeconomics and Finance: Theory and Econometric Methods*. Blackwell, Oxford, UK.

Marteau, R., Sultan, B., Moron, V., Alhassane, A., Baron, C., Traoré, S.B. (2011) 'The onset of the rainy season and farmers' sowing strategy for pearl millet cultivation in southwest Nigeria'. *Agricultural and Forest Meteorology* [Online]. 151 (10): 1356–1369.

Mutuc, M., Pan, S., and Hudson, D. (2010) 'Response of Cotton to Oil Price Shocks'. Available from:
<https://ageconsearch.umn.edu/bitstream/56425/2/Response%20of%20Cotton%20to%20Oil%20Price%20Shocks.pdf> [Accessed on 15th March 2018]

Nazlioglu, S., and Soytas, U. (2012) 'Oil price, agricultural commodity prices, and the dollar: a panel cointegration and causality analysis'. *Energy Economics*, 34: 1098–1104.
Obstfeld, M., and Rogoff, S. (1996) *Foundations of International Macroeconomics*. Cambridge, MA: MIT Press.

Osborne, J. & Waters, E. (2002) 'Four assumptions of multiple regression that researchers should always test'. *Practical Assessment, Research & Valuation* [Online]. 8(2).

Pindyck, R., and Rotemberg, J. (1990) 'The Excess Co-Movement of Commodity Prices'. *The Economic Journal* [Online]. 100 (403): 1173-1189.

Plastina, A. (2010) 'Hedge Funds and Cotton Prices'. *Cotton: Review of the World Situation* [Online]. 63(3).

Reddy, K., Hodges, H., McKinion, J. (1995) 'Carbon dioxide and temperature effects on pima cotton growth'. *Agriculture, Ecosystems and Environment* [Online]. 54: 17–29.

Reddy, K.R., Koti, S., Davidonis, G.H., Reddy, V.R. (2004) 'Interactive effects of carbon dioxide and nitrogen nutrition on cotton growth development, yield, and fiber quality'. *Agronomy Journal* [Online]. 96: 1148–1157.

Roth, G., Harris, G., Gillies, M., Montgomery, J., Wigginton, D. (2013) 'Water-use efficiency and productivity trends in Australian irrigated cotton: a review'. *Crop & Pasture Science*, 64: 1033–1048.

Saghaian, H. (2010) 'The impact of the oil sector on commodity prices: correlation or causation?'. *Journal of Agricultural and Applied Economics* [Online]. 42: 477–485.

Silvennoinen, A., and Thorp, S. (2010) 'Financialization, Crisis and Commodity Correlation Dynamics'. Research Paper Series No. 267, Quantitative Finance Research Centre, University of Technology, Sydney.

Traore, B., van Wijk, M.T., Descheemaeker, K., Corbeels, M., Rufino, M.C., Giller, K.E., (2014) 'Evaluation of climate adaptation options for Sudano-Sahelian cropping systems'. *Field Crop Research* [Online]. 156 (0), 63–75.

Trenberth, K.E., & Stepaniak, D.P., & Caron, J.M. (2000) 'The global monsoon as seen through the divergent atmospheric circulation'. *Journal of Climate* [Online]. 13: 3969–3993.

USITC (2011) 'China's Agricultural Trade: Competitive Conditions and Effects on U.S. Exports'. USITC Publication no. 4219.

Vogel, F. A. & Bange, G. A. (1999) 'Understanding USDA crop forecasts'. U.S. Department of Agriculture, National Agricultural Statistics Service and Office of the Chief Economist, World Agricultural Outlook Board (1554).

Westfall, H. (2014) 'Kurtosis as Peakedness, 1905–2014. R.I.P.'. *The American Statistician* [Online]. 68(3): 191-195.

WFE (2013). WFE Database. Available from: <http://www.world-exchanges.org> [Accessed on 15th March 2018]

Williams, A., White, N., Mushtaq, S., Cockfield, G., Power, B., Kouadio, L. (2015) 'Quantifying the response of cotton production in eastern Australia to climate change'. *Clim. Change* [Online]. 129: 183–196.

Wooldridge J. (2012) *Introductory Econometrics: A modern approach* (5th edition). USA: Cengage Learning.

World cotton production bales (2018). Available from: <http://www.cottoninc.com/corporate/Market-Data/MonthlyEconomicLetter/pdfs/English-pdf-charts-and-tables/World-Cotton-Production-Bales.pdf> [Accessed on 15th March 2018]

Yu, E., and Bessler, A., and Fuller, S. (2006) 'Cointegration and Causality Analysis of World Vegetable Oil and Crude Oil Prices'. In: *The American Agricultural Economics Association Annual Meeting*; Long Beach, California. 23–26 July.

Zhang, H., and Dufour, J., and Galbraith, J. (2016) 'Exchange rates and commodity prices: Measuring causality at multiple horizons'. *Journal of Empirical Finance* [Online]. 36:100–120.

Zhang, Q., and Reed, M. (2008) 'Examining the Impact of the World Crude Oil Price on China's Agricultural Commodity Prices: The Case of Corn, Soybean, and Pork.' In: *The Southern Agricultural Economics Association Annual Meetings*; Dallas, TX: 25 February.

Zuidema, P. & Fairall, C. (2007) 'On air-sea interaction at the mouth of the Gulf of California'. *Journal of Climate* [Online]. 20(9).

