Sveriges lantbruksuniversitet
Swedish University of Agricultural Sciences

**Department of Economics**

# Volatility of copper prices and the effect of real interest rate changes

– Does the theory of storage explain the volatility of copper spot and futures prices?

*Moa Duvhammar*

Master's thesis • 30 hec • Advanced level
Environmental Economics and Management – Master's programme
Degree project (A2E) • No 1131 • ISSN 1401-4084
Uppsala 2018

**Volatility of copper prices and the effect of real interest rate changes**

**-** Does the theory of storage explain the volatility of copper spot and future prices?

*Moa Duvhammar*

**Supervisor:** Yves Surry, Swedish University of Agricultural Sciences, Department of Economics

**Examiner:** Rob Hart, Swedish University of Agricultural Sciences, Department of Economics

**Credits:** 30 hec
**Level:** A2E
**Course title:** Independent project in economics
**Course code:** EX0811
**Programme/education:** Environmental Economics and Management – Master's programme
**Faculty:** Faculty of Natural resources and Agricultural Sciences

**Place of publication:** Uppsala
**Year of publication:** 2018
**Cover picture:** Pixabay
**Title of series:** Degree project/SLU, Department of Economics
**Part number:** 1131
**ISSN:** 1401-4084
**Online publication:** http://stud.epsilon.slu.se/

**Keywords:** Theory of Storage, Copper Price Volatility, Futures curve, Conditional Variance, GARCH

**Sveriges lantbruksuniversitet**
**Swedish University of Agricultural Sciences**

Department of Economics

## Abstract

The purpose of this thesis is to determine if the predictions of the theory of storage can explain the volatility of copper prices during the past two decades. The theory predicts that decreasing interest rates should reduce the volatility of commodity prices by encouraging the smoothing of short-run price swings caused by temporary shocks to supply and demand. In contrast, interest rates should have no effect on price volatility in the long-run as inventory smoothing cannot be used against persistent shocks. The theory is tested by estimating the volatility of copper spot and futures prices traded on the London Metal Exchange (LME) using the Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model for the period of 1994 to mid-2017. The effect of real interest rates changes on the volatility of the prices is also examined. Temporary shocks are identified by movements in the time spread of the futures curve, calculated as the price difference between the 15-months' contract and the spot contract. The volatility effect of persistent shocks is represented by fluctuations in long-term prices in terms of the 15-months' and 27-months' contracts.

The empirical results show that the volatility of copper prices have been largely driven by persistent shocks during the sample period and that the real interest rate has a significant decreasing effect on the volatility of all contracts, including long-term prices. This suggest that if the expectations of booming demand for copper and increasing interest rates are realized in the coming years, the volatility of copper is likely to increase considerably. This will have important implications to a number of countries and industries, such as the growing sectors of renewable energy systems and technologies which rely heavily on copper.

## Table of contents

# 1. Introduction

## 1.1. Background

Copper prices have become increasingly volatile since the beginning of the 2000s. This is illustrated in Figure 1, displaying a significant increase in price fluctuations since 2003. Commodity prices have historically undergone periods of boom and bust, entailing long periods of deviations from the long-run trend. These cycles, sometimes decades-long, have been associated with persistent demand shocks driven by world GDP (Jacks, 2013; Stürmer, 2016), and have eventually been punctuated by booms and busts in prices. The booms and busts have become increasingly longer and larger in more recent years and are particularly bearing in determining the volatility of real commodity prices (Jacks, 2013).


Figure 1: Spot price of copper, February 1994 - September 2017

As for the factors influencing short-term fluctuations, macroeconomic forces such as real interest rates (e.g. Frankel and Hardouvelis, 1985; Barsky and Kilian, 2002; Hamilton, 2008; Frankel, 2008) and exchange rates (Akram, 2009) have been identified as drivers of commodity prices along with changes in inflation, industrial production, inventories and the long-term and short-term interest rate spread, particularly during periods of high volatility (Karali and Power, 2013). Focusing on real interest rates, high real interest rates affect real commodity prices through three channels (Frankel, 2008): (i) by decreasing the firm's demand for inventories, as the interest rate constitutes a financial cost of storage (ii) by increasing the incentive to extract the commodity today rather than later and earn interest on the proceeds from the sale, and (iii) by encouraging speculators to shift out of commodity contracts, especially spot contracts, and into treasury bills. All three mechanisms work to reduce the market price of commodities as market supply increases, while a decrease in the real interest rate has the opposite effect. An issue with examining the causality of interest rates on commodity prices is however that they both are affected by the business cycle. Akram (2009) addresses this by controlling for factors relating to economic growth and find that commodity prices rise when the real interest rate fall and when the real value of the dollar depreciates. Furthermore, oil and metal prices show overshooting behaviour in response to interest rate changes such that current prices rise more than the long-run equilibrium level.

As a result of the inherent volatility of commodity prices, market participants have always sought ways of hedging against price fluctuations. Futures contracts are among the most popular financial instruments for managing risk and have been trading for hundreds of years. It was however only in the beginning of the 2000s that commodity futures became popular in mainstream investment portfolios. Greer (2000) could demonstrate a negative correlation between commodity returns and stock returns and after the equity market crash in 2000, billions of dollars flowed into commodity markets from a range of financial institutions, such as hedge funds, insurance companies and pension funds. The increase in financial speculation on commodity futures thus occurred at the same time as commodity prices became increasingly volatile. Tang and Xiong (2012) argue that this increase in speculation on commodity markets has made commodity prices increasingly correlated with one another and with the stock market, which should explain the increased volatility of non-energy commodities that occurred around 2008. Prior to the beginning of the 2000s, commodity prices were largely uncorrelated with one another (Erb and Harvey, 2006) or with the stock market (Gorton and Rouwenhorst, 2006) and individual commodity prices were largely determined by supply and demand factors. The increased correlation across commodities and with the stock market has, however, exposed commodity prices to the general risk appetite for financial assets and the investment behaviours of commodity index investors (Tang and Xiong, 2012). Contrary to these explanations, Gruber and Vigfusson (2016) found that the correlation and volatility of commodity prices have increased due to decreasing interest rates and increasing volatility of persistent shocks. Their results are in line with the theory of storage, which is well established in the literature and form the conceptual framework of this thesis.

## 1.2. The problem

The theory of storage predicts that the volatility of real commodity prices should fall with decreasing real interest rates. The rationale behind this is that interest rates constitute a cost of storage for consumers and producers, and reduced storage costs (lower interest rates) should therefore encourage the use of inventories to smooth price fluctuations originating from temporary shocks. Inventory smoothing implies that, in a situation of e.g. a temporary spike in demand accompanied by higher prices, producers can sell out of inventories and profit on the temporary "shock" in prices. This is of course conditional on the level of inventories that have been carried into the current period. If interest rates have been high, storage has been costly and inventory levels are likely to be low, which would make the market more vulnerable to unexpected shocks to supply and demand. In contrast, price fluctuations originating from persistent, long-lasting shocks do not encourage inventory smoothing as it is not profitable in the long-term and inventories would eventually be depleted.

The theory thus predicts that periods of low interest rates should display lower volatility in commodity prices. The increased volatility in copper spot prices plotted in Figure 1 occurred during a period of relatively low interest rates. According to the theory and the empirical findings of Gruber and Vigfusson (2016), this is explained by an increase in the volatility of persistent shocks against which low interest rates have no impact. Their empirical results show however no significant impact of the real interest rate on the volatility of copper prices, and the same was found by Hammoudeh and Yuan (2008). Several studies have tested the theory of storage in terms of the relationship between inventory levels and copper price volatility (e.g. Fama and French, 1988; Ng. and Pirrong, 1994; Brunetti and Gilbert; 1995; Geman and Smith, 2012) and found evidence supporting the theory but the prediction relating decreasing real interest rates with lower copper price volatility has been less explored.

Addressing this gap in the literature is important for several reasons. Copper is one of the most important metals in the world by production volumes and variety of applications, providing crucial materials to buildings, infrastructure, power lines, and electronics. It is the main source of export earnings for producing countries like Chile, Peru and Zambia, generating important employment, investments and government revenue. In addition, demand is expected to rise significantly in the years ahead due to the increasing global population and economy, and the growing sectors of renewable energy systems and technologies. Copper is a crucial input component in electric vehicles and efficient motors, as well as wind turbines and solar panels and the infrastructure that powers them. Volatility in copper prices thus constitutes a risk factor affecting long-term investment decisions concerning these new important systems and technologies, which are largely driven by decreasing costs. Price volatility also has a negative impact on macroeconomic factors, particularly for producing countries and low-income countries. Exploring the characteristics of copper volatility is therefore relevant in light of the expected increases in interest rates in the near future, the booming global demand for copper, and the importance of the metal for a wide range of industries and countries.

### 1.3. Objective of the thesis

The purpose of the study is to determine if the volatility of real copper prices for the past two decades can be explained by the theory of storage, which predicts that decreasing real interest rates should have an important dampening effect on commodity price volatility. To test if the predictions of the theory hold, the volatility of monthly copper spot and futures prices and the volatility effect of real interest rate changes will be estimated using the Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model. The research question to be tested is:

- *Are the characteristics of copper price volatility consistent with the predictions of the theory of storage?*

The hypotheses derived from the theory are the following:

- *Hypothesis 1:* Short-term prices are more volatile than long-term prices as they are affected by both temporary and persistent shocks.
- *Hypothesis 2:* Decreasing real interest rates reduce the volatility of short-term prices as well as the spread between futures and spot prices (hereafter termed the "time spread").
- *Hypothesis 3:* Decreasing real interest rates have no effect on the volatility of long-term prices.

### 1.4. Outline

The rest of the thesis is organized as follows. Section 2 provides the relevant theoretical framework of the thesis, including a brief introduction to the market for copper futures followed by a literature review of the theory of storage. Section 3 describes the empirical methodology used throughout the thesis by reviewing the features of time series data and the conditions that must be satisfied before estimating the GARCH models. Section 4 presents the data, relevant diagnostic tests on the variables, and specifies the models to be estimated. Section 5 presents the empirical results, which are discussed in Section 6. Finally, Section 7 summarizes and concludes the thesis.

# 2. Theoretical framework

## 2.1. The futures market for copper

As a consequence of the inherent volatility of commodity markets, producers and consumers have always sought ways of hedging and trading the risk of large price fluctuations. This resulted in the development of commodity futures markets in which options and futures contracts are traded (Geman and Smith, 2013). A futures contract allows participants to lock in a price in advance and obliges the owner to pay to the seller on the maturity date and in return receives a specified quantity of the commodity. The maturity typically ranges from one month up to several years in the future. For example, a construction company may want to fix the price of copper that they will use some months later to avoid unexpected price increases. Spot markets also exist, in which immediate delivery is available (or typically in two days forward). The copper spot price is the actual price paid when for example large manufacturers buy the quantity they need for the production.

The standard method of estimating the volatility of commodity prices is to use data on futures prices. Futures prices are available at high sampling frequency and the contracts are standardized such that e.g. the quality is the same across prices. In addition, price discovery usually occurs in futures markets (Karali and Power, 2013). This implies that actual commodity prices are determined in the futures market based on supply and demand factors related to the market. Futures prices thus reflect the expected spot price at a future date and vary depending on market expectations on scarcity, extraction costs and inventory levels. Futures and spot prices are thus interlinked and can be used to estimate the volatility of real copper prices over time.

Copper prices are suitable for analysis for several reasons. First of all, it is a relatively homogenous good as refined copper is 99.99 per cent pure copper. Secondly, the cost of transporting copper constitutes a small percentage of the final price. Third, the supply of copper is subject to little seasonal variation and only minor in demand (related to slight variations in construction activity across the northern hemisphere year) and it is also easy to store at a relatively low cost and with negligible degradation over time compared to other commodities (Geman and Smith, 2013). Finally, copper is sold on global markets rather than in various regional markets and as a result, the prices are correlated within the bounds set by the cost of transporting copper (Svedberg and Tilton, 2006). The London Metal Exchange (LME) was founded in 1877 and is the futures exchange with the world's largest market in options and futures contracts on metals and provides 600 warehouses worldwide. It is the principal marketplace to establish prices in the copper market (Stürmer, 2016).
On each trading day, contracts for delivery in 2 days ("spot"), 3 months, 15 months and 27 months are traded. The 3-month contract is the most traded contract and was originally introduced because it took that long for tin from South-East Asia, or copper from Chile, to arrive by ship to London (Geman and Smith, 2013).

## 2.2. Theory of storage

The theory of storage has become the dominating theory explaining short-term fluctuations in futures and spot prices. In its simplest form, the framework takes the supply of the commodity as given and assumes that risk-neutral commodity consumers operate in a competitive market in which they choose the optimal quantity to consume or store, based on the price of storage. The price of storage relates to the difference between the futures price and the spot price (or the price of the contract closest to maturity if spot prices are not available). In the U.S. in the early 1930s, empirical research had long noted that short-term futures prices were often higher than long-term futures, reflecting a negative spread (Keynes, 1930). This was a puzzle to researchers at the time, as inventories were carried over periods despite market expectations of decreasing prices. The rational strategy would be to buy the commodity later at a lower price, or buy the long-term futures contract, but market inventory levels were nevertheless substantial despite the apparent capital loss. In addition, the spread between futures and spot prices also seemed to vary from year to year. Working (1933, 1934, 1948, and 1949) sought to explain this by analysing the futures market on wheat and inventory levels. His findings laid the foundation to the theory of storage, relating spot and futures prices with market inventory levels. Working plotted the futures—spot spreads observed over time against observed inventory levels and a clear relationship emerged: years of low inventory displayed higher spot prices than futures prices, resulting in a negative spread, while years of no shortage (high inventory levels) displayed futures prices slightly above spot prices, which approximately corresponded to the cost of storing the commodity until the future delivery date. The relationship has been termed the Working curve (1933) and is shown in Figure 2. The curve shows a negative spread at times of low inventory levels and a slightly positive spread, approaching a constant level, for high levels of inventories.

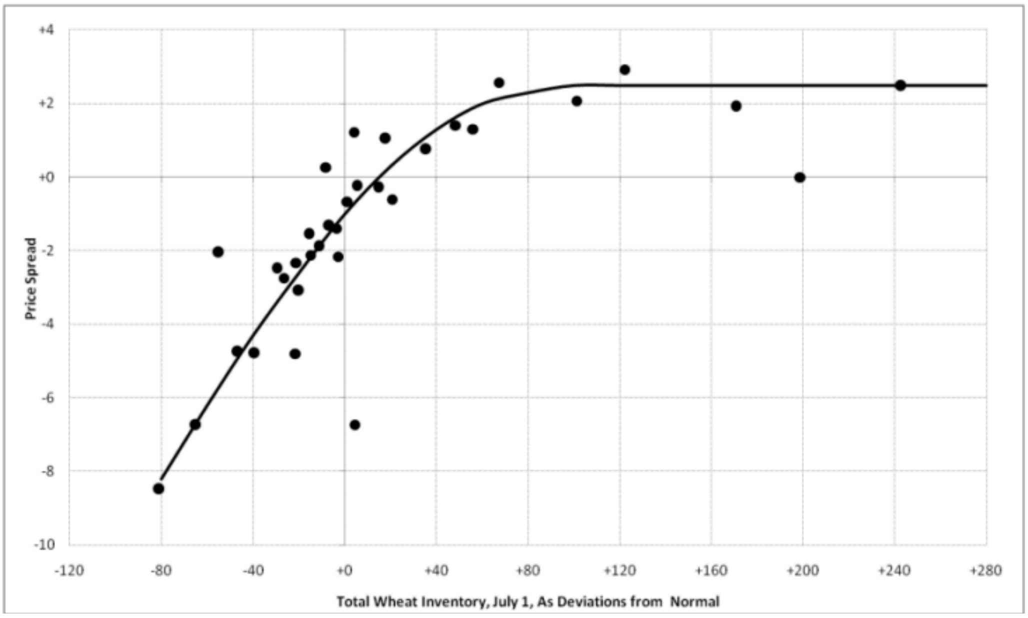



Figure 2: Relationship between Chicaco July-September spread in June and U.S. wheat stocks on July 1 (Working, 1933)

Working also contributed importantly to demonstrating some typical features of commodity futures prices which are in line with the predictions that will be tested later in the thesis, summarized as follows:

(i) Spot price volatility is higher in times of low inventories compared to times of high levels since any news on short term supply, demand or inventory will have a large impact on the spot market.
(ii) Information concerning the short-term supply or demand affects short-term prices more than the long-term futures.
(iii) Information affecting long-term supply or demand affects the short and long-term futures prices approximately equally, thus leaving the spread unaffected.

The conflict of holding inventories in times of negative futures—spot spreads was however not solved until Kaldor (1939) provided an explanation by introducing the concept of an unobserved "convenience yield". The convenience yield represents the benefit derived from holding physical commodities and having ready access to them, as it allows firms to immediately respond to demand and supply shocks. The convenience yield enters the relationship between the futures price and the spot price as follows (Fama and French, 1987):

$$F(t,T) - S(t) = S(t)[R(t,T) + C(t,T) - Y(t,T)] \tag{A1}$$

Where,

- $F(t,T)$ is the futures price of a commodity at time $t$, for delivery at time $T$.
- $S(t)$ is the spot price at time $t$.
- $F(t,T) - S(t)$ is the return of buying the commodity in time $t$ and selling it for delivery at time $T$.
- $\mathrm{R(t,T)}$ is the cost of financing the futures position from time $t$ to time $T$, in other words the interest rate.
- $C(t,T)$ is the cost of storage of the physical commodity, such as warehouse costs, from time $t$ to time $T$.
- $\mathrm{Y(t,T)}$ is the convenience yield associated with storing the commodity from time $t$ to time $T$, calculated to satisfy the relationship rather than observed directly.

The theory predicts a negative relationship between the convenience yield and inventories, implying that the smaller the level of inventories on hand the greater the convenience yield of an additional unit of storage (Brennan, 1958). This implies that the marginal convenience yield can sometimes exceed the marginal costs of storage when inventory levels are low, thus resulting in the negative futures—spot spreads observed in Working's curve in Figure 2. The spread can therefore be used as a signal for the market level of inventories: a positive spread signals a soft market with inventories in abundance, and a negative spread signals a tight market with inventories running low (Frankel, 2014). To better understand the relationship between the spread and the convenience yield, equation (A1) can be expressed as:

$$\frac{(F(t,T) - S(t))}{S(t)} = R(t,T) + C(t,T) - Y(t,T) \tag{A2}$$

The relationship can also be expressed in terms of the convenience yield, which gives:

$$\frac{(F(t,T) - S(t))}{S(t)} - R(t,T) - C(t,T) = -Y(t,T) \tag{A3}$$

From (A3), it is clear that the convenience yield is simply the futures—spot spread but expressed with opposite sign and adjusting for the cost of financing and storing the commodity over the period (Geman and Smith, 2013).

Working's finding of high spot price volatility in times of low inventory levels has been examined in several studies related to base metals. Fama and French (1988) examined the relationship between the volatility of base metals traded on the LME and inventories and found that, in line with Working's prediction, spot price volatility increases as inventory decrease. This relationship is supported by Ng. and Pirrong (1994) who found a strong relationship between the spread and spot price volatility, followed by Brunetti and Gilbert (1995), who also linked high spot price volatility of LME-traded base metals to low inventory levels.

Deaton and Laroque (1992, 1996) extended the model and confronted the theory by testing whether it could explain the actual behaviour of a broad range of commodity prices. The authors failed to reproduce the predictions of the theory, such as the high autocorrelation of most commodity prices. Cafiero et. al. (2011) pointed out problems in their estimations and when re-estimating the model, it actually yielded estimates consistent with observed levels of autocorrelation. The model of Deaton and Laroque entails forward-looking stockholders of a commodity, who maximizes profit by considering the expected future price relative to the current price as well as the cost of carrying inventories into the next period. The equilibrium price of the commodity, which would otherwise be determined by a simple process of supply and demand, is thus determined by the maximization process of the stockholders. The source of volatility in the model is unexpected temporary shocks to supply or demand, which can be dampened by holding inventories. In the model, commodity prices are denoted by:

$$P_t = P(z_t, \Delta I_t), \tag{A4}$$

where $P_t$ is the price, $z_t$ is a combination of supply and demand for the commodity (i.e. net demand or net supply) and subject to stochastic shocks, and $\Delta I_t$ is the inventory level for each period. Inventories are accumulated according to:

$$\Delta I_t = I_t - (1 - \partial) I_{t-1}, \tag{A5}$$

where $(I_t)$ is the inventory level in time $t$, $(1 - \partial)$ is the depreciation rate at which inventories deteriorates over the period, and $(I_{t-1})$ is the quantity of inventories carried from time $t - 1$ to time $t$. Stockholders are assumed to be profit-maximizing, risk-neutral, and hold a non-negative quantity of inventories (as commodities cannot be consumed before they exist). Risk-neutrality ensures that the futures price equals the expected future spot price. In each period, supply can either be consumed or entered into storage for future consumption and consumption can either come from inventory or current supply. Inventories are associated with costs in terms of the depreciation rate (δ) and a constant real interest rate (r), which affect the valuation of the expected price in the next period. In equilibrium, prices must thus satisfy the following relationship:

$$p_t = \max\left[\frac{(1-\delta)}{(1+r)} E_t p_{t+1}, P(z_t + (1-\delta) I_{t-1}\right] \tag{A6}$$

The first term on the right-hand side represents the expected value of storing one unit of the commodity over the period, adjusted for depreciation and interest costs. If prices are expected to increase or decrease by less than the storage cost, inventories will be zero. For inventories to be carried to the next period, the expected price must be the current price plus the storage cost, implying a positive futures—spot spread. The second term is the value of the current price if no inventories are carried into the next period, with the current net supply $z_t$ and surviving inventories from the previous period $(1 - \delta) I_{t-1}$ sold to the highest bidder. If this price is higher than the first term, speculators will be reluctant to hold inventories and the latter price will set the market price. However, if selling everything (the value of the second term) would drive the price lower than the expected price net of costs, inventories would be held until the price equals the first term in the brackets and arbitrage was no longer profitable. In other words, even though the interest rate is assumed to be constant, the model suggests that lower interest rates decrease price fluctuations in response to temporary shocks by encouraging stockholding. This prediction will be tested in the empirical results section.

# 3. Empirical methodology

## 3.1. Stationarity

A time series is a sample of random variables ordered in time, often called a stochastic process, which can either be stationary or non-stationary. The definition of a strictly stationary process is when the joint probability distribution is not dependent on time, and the mean and variance characterizing the distribution are stable over time (Maddala and Kim, 1998). Regressing a non-stationary variable on another could provide statistically significant results when in fact there is no causal relationship between the variables. This is called a spurious regression and was coined by Granger and Newbold in 1974. It is thus important to ensure that the variables to be modelled are stationary. Sources to non-stationarity could be trends, which means that the variable contain a unit root, or structural breaks causing a sudden change in the time series that must be accounted for in the regression model (Stock and Watson, 2012). In practice, time series variables are rarely stationary but it can be achieved through differencing the variable so that trends or other factors causing the non-stationarity are removed (Maddala and Kim, 1998). In addition, Meucci (2005) argues that expressing prices in their logarithmic (log) first difference form, i.e. in terms of log-returns between the price of one period and another, will simplify the modelling since log-returns are approximately symmetrically distributed in contrast to linear returns.

Stationarity can be examined by plotting the time series and examining whether they appear to include trends or other systematic structures. Tests can also be applied to determine if the variable follows a unit-root process. In this study, the Augmented Dickey-Fuller (ADF) test will be applied to test for unit roots, against the alternative hypothesis that it was generated from a stationary process. Dickey and Fuller (1979) introduced the test which, in addition to testing for unit root as the original Dickey-Fuller test, controls for serial correlation by fitting a model of the form:

$$\Delta y_t = \alpha + \beta y_{t-1} + \zeta_1 \Delta y_{t-1} + \cdots + \zeta_k \Delta y_{t-k} + \epsilon_t, \tag{B1}$$

where $\Delta y_t$ is the first difference form of the dependent variable, $\alpha$ is a constant, $\beta$ is the coefficient of the autoregressive dependent variable, and $k$ is the number of lags of the autoregressive process chosen in the test. The null hypothesis of the test is that $\beta = 0$ against the alternative that $\beta < 0$. Stationarity thus implies that $y_t$ returns to a constant mean and can therefore be used to predict the next period's change while non-stationary variables are random walks, and cannot be used to forecast values of the consecutive periods. In addition, a supremum Wald (sup Wald) test will be applied for testing if the time series variables are subject to structural breaks. The sup Wald test computes sample statistics over a set of possible break dates for a range of the data. Andrews (1993) recommends trimming the sample to be tested by 15% so that, for this study, observations during October 1997 to March 2014 will be tested for breaks.

## 3.2. Volatility clustering

Another feature of time series data, particularly for economic and financial data, is that the variance tends to be grouped in clusters (Zivot, 2008). This implies that periods subject to particularly large shocks or disturbances are followed by large variances in consecutive periods and vice versa for small variance. Volatility clustering usually implies that the variance of the error term of the regression is higher for some ranges of the data than for others and that the change in the variance is not random (Stock and Watson, 2012). Instead, the variance of the error term is likely to be correlated with past values and thus suffer from time-varying heteroskedasticity, or conditional heteroskedasticity. This clustering of the variance of the error term over time violates the ordinary least squares (OLS) assumption of homoscedasticity, which states that the expected value of the sum of squared errors does not vary over time (Engle, 2001). OLS can therefore not be used as estimation method since it would produce biased regression coefficients. Conditional heteroskedasticity is, however, not a problem one need to correct for if the key issue is to analyse why the variance of the error terms changes. There are models designed to capture conditional heteroskedasticity in the regression error, namely the autoregressive conditional heteroskedasticity (ARCH) model and its extensions.

### 3.3. ARCH and GARCH models

The ARCH model was first introduced by Engle in 1982 to model volatility when the variance of the error term varies over time. Today, estimating and forecasting volatility is a central part of financial econometrics and the ARCH model and its extensions have become the standard tool for doing it. The ARCH model is a system of two equations: the conditional mean equation and the conditional variance equation, where conditional implies that the mean and variance are time-varying such that they are conditional on the information set available up to time $t-1$.
Consider the following conditional mean model of price returns at time $t$ $(R_t)$ (Stock and Watson, 2012):

$$R_t = \mu + \eta' X_{t-i} + \varepsilon_t, \tag{B2}$$

where $R_t$ is a function of a constant $(\mu)$ denoting the average returns (price difference of two periods), a set of variables $(X_{t-i})$ that could be autoregressive terms of $R_t$ and other exogenous variables affecting the returns, $(\eta)$ are the associated coefficients to be estimated, and $(\varepsilon_t)$ the error term. In Engle's original model (1982), the error term$(\varepsilon_t)$ was assumed to follow a normal distribution with zero mean and a variance of $\sigma_t^2$ but, as pointed out by Mandelbrot (1963) and many others, the distribution of financial time series is often leptokurtic. A leptokurtic distribution implies that the series have a higher peak and heavier tails than a normally distributed sample, implying that extreme values are more frequent than would be expected with a normal distribution. It is thus common to fit the model assuming the errors to follow distributions with fatter tails than the normal distributions, such as the Student t-distribution or the generalized error distribution (GED) (Zivot, 2009).

The ARCH (q) model estimates the conditional variance and is specified as:

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 \varepsilon_{t-2}^2 + \cdots + \alpha_p \varepsilon_{t-q}^2, \tag{B3}$$

where the conditional variance $(\sigma_t^2)$ at time $t$ is estimated as a function of past squared values of $\varepsilon_t$ and unknown parameters $(\alpha_0, \alpha_1, \ldots, \alpha_p)$ to be estimated. This way, the magnitudes of the parameters $(\alpha_i)$ indicate the importance of past values for the current volatility. The econometric challenge is to determine the order of $q$ and in general, it requires a large number of lags to capture the effect of volatility clustering. This can make ARCH estimations complicated.

Bollerslev (1986) extended the model such that $\sigma_t^2$ depends on its own lags in addition to the lags of the squared errors. The model is called the Generalized ARCH (GARCH) model and has fewer parameters than the ARCH and is thus easier to estimate. Hansen and Lunde (2004) have also shown that it is difficult to find a volatility model that beats the simple GARCH (1, 1) model, including one lag of the error term and the variance.

The conditional mean equation is the same as for the ARCH model but the conditional variance equation, the GARCH (p, q), is specified as:

$$\sigma_t^2 = \omega + \alpha_1\varepsilon_{t-1}^2 + \cdots + \alpha_q\varepsilon_{t-q}^2 + \beta_1\sigma_{t-1}^2 + \cdots + \beta_p\sigma_{t-p}^2 + \delta y_t, \tag{B4}$$

where the conditional variance ($\sigma_t^2$) at time t depends on a constant (ω), the effect of past squared errors $\alpha_q \varepsilon_{t-q}^2$, the effect of past values of the variance itself $\beta_p\sigma_{t-p}^2$, and the effect of some exogenous variable $\delta y_t$. The parameter $\alpha_p$ is also called the ARCH effect, or past shock effect, and $\beta_p$ is called the GARCH effect and represents the past volatility effect.
Since the variance is a function of the mean, equation (B2) and (B3 or B4) are estimated simultaneously using the maximum likelihood method, which maximizes the log likelihood function with respect to the parameters ω, $\alpha_q$, $\beta_p$ and $\delta$. The sum of $\sum(\alpha + \beta)$ measures the degree of convergence to the long-run equilibrium level of the variance, or in other words the persistence of random shocks to the conditional variance in the model (Ma et al., 2006). A high value of the sum indicates slow convergence, or high persistence. To ensure a positive conditional variance, the estimated parameters of the variance equation must be non-negative, such that $\omega \geq 0$, $\alpha \geq 0$ and $\beta \geq 0$ and $|\alpha + \beta| < 1$.

### 3.4. EGARCH model

Nelson (1991) was first to recognize the non-negativity restriction on the parameters (ω, $\alpha$ and $\beta$) as a weakness of the GARCH model. As the residuals in equation (B4) enter as squared errors, the effect of past residuals or shocks can only be symmetric on to the conditional variance. As a result, negative and positive shocks can thus not have different impacts on the variance. It is a stylized fact that negative shocks in terms of bad news have a larger impact on the volatility of stock prices than good news (i.e. positive shocks) (Zivot, 2009). For example, volatility tend to be higher in a declining market than in a rising market due to the decreasing effect of bad news on stock prices which increases the debt-equity ratio for companies making the stock more volatile. Nelson (1991) addressed this by relaxing the restrictions on the parameters by specifying the Exponential GARCH (EGARCH) model. The EGARCH uses the same mean equation as the original GARCH model but the variance equation allows for asymmetric effects on the conditional variance and is specified as:

$$\ln(\sigma_t^2) = \omega + \gamma_1\frac{\varepsilon_{t-i}}{\sigma_{t-j}} + \gamma_2\left|\frac{\varepsilon_{t-i}}{\sigma_{t-j}}\right| + \beta\ln(\sigma_{t-j}^2) + \delta y_t, \tag{B5}$$

where the conditional variance $\sigma_t^2$ is estimated in a log-linear form to allow for positive and negative impacts, ω is the intercept, and $\sigma_{t-j}^2$ the logged GARCH term and $\beta$ is its coefficient. The asymmetric effect is measured by ($\gamma_1$) and the symmetric effect (replacing the ARCH effect in the original GARCH model) by ($\gamma_2$). If $\frac{\varepsilon_{t-i}}{\sigma_{t-j}}$ is positive, then the effect of the shock on the log conditional variance is ($\gamma_1 + \gamma_2$) and if the term is negative, the effect is ($\gamma_2 - \gamma_1$).

### 3.5. Engle's ARCH test

If the squared residuals of the estimated conditional mean equation display autocorrelation, they are said to exhibit autoregressive conditional heteroskedastic (ARCH) effects. This can be determined by examining the autocorrelation function of the squared error term of the mean regression or by applying the Engle's ARCH test. The null hypothesis of the test is that the squared residuals are determined by a white noise process while the alternative hypothesis is that they are correlated with its lagged terms. The alternative hypothesis is thus that the squared error term ($\varepsilon_t$) in equation (5) is correlated according to:

$$\varepsilon_t{}^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}{}^2 + \cdots + \alpha_m \varepsilon_{t-m}{}^2 + u_t, \tag{B6}$$

where $u_t$ is a white noise error process. The null hypothesis of the test is that all the coefficients are zero, such that:

$$\alpha_0 = \alpha_1 = \cdots = \alpha_m = 0, \tag{B7}$$

and no ARCH effects are present since the error is only dependent on $u_t$. The test will be applied to the residuals of the mean models in the study.

# 4. Data and descriptive statistics

According to the predictions of Deaton and Laroque's model, decreasing interest rates should reduce volatility in prices attributable to temporary shocks. In contrast, decreasing interest rates should not affect price volatility caused by persistent shocks. This study will follow the method of Gruber and Vigfusson (2016), first introduced by Schwartz and Smith (2000), and test the theory of storage by separating price movements caused by temporary shocks from persistent shocks.

## 4.1. Data

To estimate copper price volatility, monthly data on spot and futures prices will be analysed. The data set covers the period of February 1994 to September 2017 and has been obtained from Thomson Reuters Datastream. The LME is the principal marketplace to establish prices in the global copper market (Stürmer, 2016). Copper prices are therefore represented by prices of the cash contract, the 3-months' futures contract, the 15-months' futures contract, and the 27-months' futures contract, all traded on the LME. The cash price is the current price of cash LME contracts for delivery two days forward. This implies that the cash price is the price paid for copper on the spot market and thus represents the spot price. All prices are expressed in U.S. dollar (USD) per metric ton. The USD currency is used since most contracts in international commodity trade are settled in USD (Kornher and Kalkuhl, 2013). As the purpose is to test the theory of storage, in which decisions are made based on relative prices over periods, interest rates and copper prices will be deflated by the U.S. consumer price index (CPI of all items) and expressed in real terms. Real prices have been calculated by taking the log of the price for each month and subtract the log of the U.S. CPI as of the same month.

The interest rate is represented by the U.S. interest rate in terms of the 3-month Treasury bill: secondary market rate. The real interest rate has been calculated by subtracting off the change in the CPI between month $t$ and $t - 1$. The U.S. rate has been chosen since global real rates tend to follow the same path as the rate in the U.S. (Gruber and Vigfusson, 2016). Finally, the time spread is calculated as the difference between the real log-price of the 15-months' contract and the real log-price of the spot price, defined as:

$$\ln\left(\frac{F_{t+15}}{S_t}\right), \tag{C1}$$

where $F_{t+15}$ is the futures price of the 15 months' contract and $S_t$ is the spot price. Illustrated in the formula, movements in the spread can only originate from temporary shocks as persistent shocks would impact both the numerator and the denominator.

## 4.2. Descriptive statistics

The descriptive summary statistics for the variables of interest are reported in Table 1, including the log-prices of the different contracts and the first differences (referred to as log-returns). The prices have close to the same mean, although the mean prices of the short-term contracts are slightly higher on average. The standard deviations (SD) of the different prices confirm the common perception that commodity prices are volatile, particularly for the shorter-term contracts, which display monthly volatilities of over 7% (for the log-returns). The 27-months' contract has the lowest average return and the lowest standard deviation.

As described in Section 3.1, log-returns have kurtosis and skewness closer to the normal distribution compared to log-prices. This is confirmed visually when inspecting the histograms in Figures 1.a-h

(Appendix 1), which show that the log-returns are more symmetrically distributed than the log-prices although the distribution is clearly not normal. In contrast, the log-returns appear to exhibit the typical distribution of financial time series data, namely asymmetry and leptokurtosis, which means that extreme log-returns are more frequent than would be expected if the returns were normally distributed.

Table 1: Descriptive statistics

| Descriptive statistics | Mean | SD | Kurtosis | Skewness |
|---|---|---|---|---|
| Log of Spot-price | 3.722 | 0.511 | 0.511 | 0.511 |
| Log of 3-month-price | 3.718 | 0.507 | 0.507 | 0.507 |
| Log of 15-month-price | 3.692 | 0.498 | 0.498 | 0.498 |
| Log of 27-month-price | 3.671 | 0.487 | 0.487 | 0.487 |
| Time spread between 15-month-price and the Spot-price | -0.030 | 0.078 | 0.078 | 0.078 |
| Real Interest Rate based on 3-month T-bill | 2.228 | 2.164 | 2.164 | 2.164 |
| Return on Spot-prices | 0.003 | 0.075 | 0.075 | 0.075 |
| Return of 3-month-prices | 0.003 | 0.072 | 0.072 | 0.072 |
| Return of 15-month prices | 0.003 | 0.064 | 0.064 | 0.064 |
| Return of 27-month prices | 0.002 | 0.060 | 0.060 | 0.060 |
| First difference of the RIR | -0.009 | 0.426 | 0.426 | 0.426 |
| First difference of the time spread | -0.000 | 0.028 | 0.028 | 0.028 |

As expected, when testing the skewness and kurtosis jointly, normality is rejected for all variables at the 1% level except from the Real Interest Rate (RIR), which is normally distributed. This suggests that assuming error distributions with fatter tails are likely to fit the data better than the normal distribution.

Moreover, looking at Figures 2.b-g) and 3.a-f) (Appendix 2 and 3), volatility clustering seems to be present based on the variance behaviour of the variables and the variance of the residuals from the mean models. Months of small volatility are followed by small volatility in consecutive months and vice versa for periods of high volatility. The period of 2006 - 2010 is marked by exceptionally large price swings.

Table 2: Period 1, February 1994 – November 2005

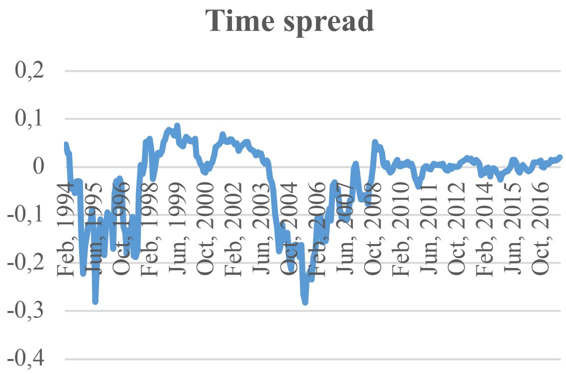
| VARIABLES | Mean | SD | Kurtosis | Skewness |
|---|---|---|---|---|
| Return on Spot-prices | 0.004 | 0.064 | 0.064 | 0.064 |
| Return of 3-month-prices | 0.004 | 0.059 | 0.059 | 0.059 |
| Return of 15-month prices | 0.002 | 0.045 | 0.045 | 0.045 |
| Return of 27-month prices | 0.001 | 0.041 | 0.041 | 0.041 |
| First difference of the time spread | -0.002 | 0.036 | 0.036 | 0.036 |

Dividing the sample into two periods to identify any pattern of change in the descriptive statistics allows for a first comparison of whether the volatility, in terms of the sample standard deviation, has changed over time. The first period ranges from 1994 to 2005 and the second period from 2005 to mid-2017. Examining the average standard deviations of the returns of the first period in Table 2, it can be noted that volatility is much lower compared to the standard deviations of the second period (Table 3) except from the volatility of the first differenced time spread, which has decreased. Volatility of the prices has increased with 33% (Spot), 42% (3-month), 73% (15-month), 83% (27-month) between the periods, whereas the spread has decreased with 53%.

Table 3: Period 2, December 2005 - September 2017

| VARIABLES | Mean | SD | Kurtosis | Skewness |
|---|---|---|---|---|
| Return on Spot-prices | 0.001 | 0.085 | 0.085 | 0.085 |
| Return of 3-month-prices | 0.001 | 0.084 | 0.084 | 0.084 |
| Return of 15-month prices | 0.002 | 0.078 | 0.078 | 0.078 |
| Return of 27-month prices | 0.003 | 0.075 | 0.075 | 0.075 |
| First difference of the time spread | 0.001 | 0.017 | 0.017 | 0.017 |

When plotting the time spread and the RIR over time in Figure 2 and 3, the pattern of the spread follows the predictions of the theory. That is, the spread appears to be less volatile at times of low interest rates, such as after the financial crisis in 2008.



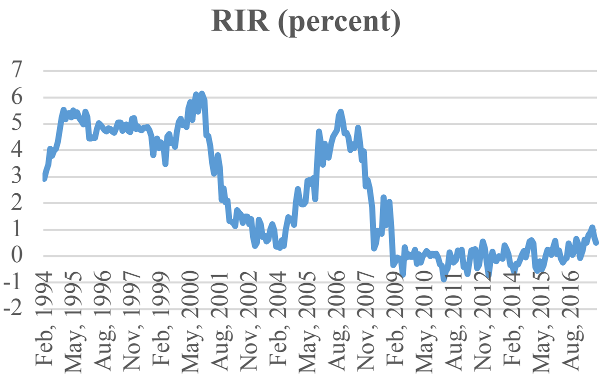




Figure 2 and 3: Time series plot of the time spread and the RIR

## 4.3. Diagnostic testing on the time series variables

Moving over to diagnostic testing of the variables and the mean models to be estimated, the log-prices are first examined in Figure 2.a) (Appendix 2). Copper prices may have grown over time and the time series plot over the sample period shows that the log-prices could include an upward trend while the time series plots of the log-returns in Figure 2.b-g) display no obvious trends. As described in section 3.1, financial time series data tend to be less variant when expressed as returns than in levels, which is also confirmed by comparing the plots of the log-prices to the log-returns. Transforming the log-prices into their first difference form and remove potential trends could thus be appropriate before estimating the models to ensure stationarity of the variables. Table 4 reports the test statistics of the Augmented Dickey Fuller (ADF) test and shows that all log-prices are non-stationary while the log-returns are stationary. The RIR is also stationary in its first-difference form but the time spread is slightly below the critical value and the null hypothesis of unit root cannot be rejected. This is also demonstrated by the autocorrelation function of the time-spread variable in Figure 4. i) (Appendix 4), which shows autocorrelation of up to 10 lags (displayed as spikes external to the confidence band). The time series of the time spread is thus likely to include a trend component. For further diagnostic testing and model estimation, all variables will therefore be modelled in their first difference form, defined as:

$$Lnr_t = lnP_t - lnP_{t-1} = \Delta lnP_t \tag{C2}$$

where $Lnr_t$ represents the first difference of the copper prices, the spread, or the RIR, defined as the difference between the log of the observation of month $t$ and $t-1$. The differenced prices will be referred to as log-returns hereafter.

Table 4: Diagnostic tests on the variables

| Diagnostic tests | Spot | 3-months' | 15-months' | 27-months' | Time-spread | First diff. Time spread | RIR | First diff. RIR |
|---|---|---|---|---|---|---|---|---|
| ADF test statistic* | -6.454 | -6.428 | -6.293 | -6.387 | -2.390 | -7.511 | -1.327 | -6.552 |
| ADF | I(0) | I(0) | I(0) | I(0) | I(1) | I(0) | I(1) | I(0) |
| Box Pierce P-values | | | | | | | | |
| Lag 1 | 0.0174 | 0.0228 | 0.0207 | 0.0823 | 0.0000 | 0.9577 | | |
| Lag 2 | 0.0202 | 0.0131 | 0.0050 | 0.0396 | 0.0000 | 0.0531 | | |
| Lag 3 | 0.0502 | 0.0338 | 0.0131 | 0.0802 | 0.0000 | 0.0469 | | |
| Lag 4 | 0.0986 | 0.0696 | 0.0281 | 0.1362 | 0.0000 | 0.0794 | | |
| Lag 5 | 0.1389 | 0.0954 | 0.0368 | 0.1480 | 0.0000 | 0.1004 | | |
| PAC | 1 | 1 | 1 | 0 | 1 | 0 | | |

*including 5 lags. Critical values for the ADF test statistic is -2.879 for the 5% significance level.

## 4.4. Diagnostic testing of the mean models

The log-return of each month is likely to be affected by the returns of previous months. To select how many lags to include in the mean models, the autocorrelation structure can be examined for each variable as well as comparing the values of the Akaike's information criterion (AIC) for each model and identify the model which minimizes the criteria. The AIC value is calculated based on the sum of squared errors of the model, the number of estimated regression coefficients, and the sample size.

Starting with the autocorrelation structure, it can be noted that the p-values of the Box Pierce Q statistics reported in Table 4 allows for rejection of the null hypothesis. The null hypothesis states that the autocorrelation up to lag *k* is zero and can be rejected for all variables at the 5% significance level up to about five lags depending on the variables examined, except for the 27-months futures price and the first difference of the time-spread. This implies that the mean models should include lagged values (except from the 27-months' contract and the first differenced spread) and to determine the lag order, the partial autocorrelation functions (PAC) for each variable can be examined, showing the unique correlation of each lag with the current value. The PACs are plotted in Figures 4. b-l) in Appendix 4, where lags external to the confidence band (the grey area) indicate the appropriate lag order for the mean models to be estimated. This can be confirmed by testing models with different number of lags and with or without the effect of the RIR. The selected mean model specifications are reported in Table 5 based on the PACs and comparing AIC values. The p-values associated with the sup-Wald tests fail to reject the null hypothesis, which means that the variables are not affected by any structural breaks except for the Time spread which, as mentioned earlier, also has a problem of non-stationarity. The spread will therefore be modelled in its first difference form.

Table 5: Mean models

| Mean model: | Spot | 3-months' | 15-months' | 27-months' | Time spread | First diff. Time spread |
|---|---|---|---|---|---|---|
| Mean specification | AR(1) | AR(1) | AR(1) | Simple | AR(1) | Simple |
| Sup-Wald test | 0.7655 | 0.7394 | 0.4943 | 0.7515 | 0.0000 | 0.8382 |
| ARCH effects p-values | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| Shapiro-Wilks p-values | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |

The third row in Table 5 shows the p-values from the Engle's ARCH test on the squared residuals from the mean models and confirms that the "no ARCH" hypothesis can be rejected for all models. The conditional variance of the variables is thus suitable to be estimated with a GARCH model. Figures 5.a-f) (Appendix 5) show the standardized normal probability plots (P-P plots) and the quantile-normal plots (Q-Q plots) of the residuals of the mean models for the different variables. The

P-P plots compares the distribution of the residuals with the normal distribution focusing on the center of the distribution, while the Q-Q plots focus on the tails of the distribution. Miller (1997) strongly recommends Q-Q plots for detecting non-normality and to reveal exactly where it differs from normality. Based on the P-P plots, the middle range of the residual distribution looks to be in line with the normal distribution except for the time spread and its first difference form, which clearly differs in the center from the normal distribution. This is likely to be due to the peakedness of the distribution, illustrated in the histogram in Figure 1.j) (Appendix 1). Based on the Q-Q plots, it can also be concluded, as discussed earlier, that the tails diverge from the normal. This is also confirmed by the Shapiro-Wilks test for normality (p-values are reported in Table 5), with the null hypothesis of normal distribution, which is rejected for all the residuals. When deciding on model specification of the models to be estimated, the normal distribution, the Student's t-distribution and the Generalized Error Distribution (GED) will be considered. The t-distribution and the GED are the most common fat-tailed error distributions for GARCH models and might thus be appropriate for the data (Zivot, 2009).

## 4.5. Model specification

Models of the GARCH family are designed to capture the different characteristics of the sample data, namely volatility clustering, non-normality of the price series, and autoregressive conditional heteroskedasticity in the errors. As described in section 3.3, the GARCH (p, q) model is a system including the regressions of the conditional mean and the conditional variance, which are estimated simultaneously. The chosen specification of the mean regression for each variable is reported in Table 5. The conditional mean of the spot contract, the 3-months' contract and the 15-months' contract will thus be estimated as first order autoregressive processes (AR(1)), defined as:

$$Lnr_t = \pi_0 + \pi_1 Lnr_{t-1} + \varepsilon_t, \tag{C3}$$

where $Lnr_t$ represents real log-returns of the copper prices defined as the log difference between month $t$ and $t-1$, $\pi_0$ is the intercept, $Lnr_{t-1}$ is the AR(1) term of $Lnr_{t,}$, and $\varepsilon_t$ is the error term. The conditional mean equation of the 27 months' contract is defined as:

$$Lnr_t = \mu + \varepsilon_t, \tag{C4}$$

where $Lnr_t$ is the real return of the 27-months' contract defined as the log difference between month $t$ and $t-1$, $\mu$ is the intercept, and $\varepsilon_t$ is the error term.
The conditional mean equation of the time spread is defined as:

$$Ln\left(\frac{F_t}{S_t}\right) - Ln\left(\frac{F_{t-1}}{S_{t-1}}\right) = \mu + \varepsilon_t, \tag{C5}$$

where $Ln\left(\frac{F_t}{S_t}\right) - Ln\left(\frac{F_{t-1}}{S_{t-1}}\right)$ is the first difference form of the time spread between the 15-months' futures price and the spot price (the cash price), $\mu$ is the intercept and $\varepsilon_t$ the error term.

To evaluate the fit of the GARCH (p, q) model for the variable of interest, a number of graphical and statistical diagnostics tests can be used. If the GARCH model is correctly specified with the appropriate order of (p, q) and underlying assumption of error distribution, the estimated standardized residuals should behave like conventional regression residuals. The standardized residuals are defined as the estimated residual of the GARCH model divided by the estimated conditional standard deviation: $\frac{\varepsilon_t}{\sigma_t} = z_t$. This implies that $z_t$ should not exhibit any autocorrelation or ARCH effects, and its distribution should approach the specified error distribution used in the estimation (Zivot, 2009).

To select appropriate GARCH (p, q) models and error distribution for the estimation of each variable, a number of different orders of the ARCH and GARCH terms have been estimated assuming the normal distribution, the Student's t-distribution and the GED distribution. As noted in Section 4.2, the log returns do not have a normal distribution due to the long tails. This could be due to the effect of past volatilities (the GARCH effect) and the normal distribution could nevertheless be adequate for the data but it is more likely that using the likelihood of the Student's t-distribution or the Generalized Error Distribution (GED), which have longer tails, will generate a better fit than the normal. This, along with the order of the lags of the GARCH and ARCH parameters (p and q), are evaluated by comparing AIC values. As expected, the Student's t-distribution gives the best fit for all models. As for the order of (p, q), the GARCH (1, 1) is selected for all variables.

Moreover, the Box-Pierce Q-statistics confirms that the squared standardized residuals (${z_t}^2$) are no longer serially correlated and the Engle's ARCH LM test fails to reject the null hypothesis of no ARCH effects. This implies that the specification of GARCH model for each variable successfully captures all ARCH effects in the errors. The conditional variance equation for all variables is specified as:

$$\sigma_t^2 = \omega + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2, \tag{C6}$$

where $\sigma_t^2$ is the conditional variance, ω is the intercept, $\alpha_1$ is the effect of past shocks to the log-return, $\varepsilon_{t-1}^2$ is the lagged value of the squared error term, and $\sigma_{t-1}^2$ represents the lagged squared conditional variance.

Continuing with the second model including the effect of the real interest rate (RIR), the GARCH (1, 1) model assuming the t-distribution is selected for all variables according to:

$$\sigma_t^2 = \omega + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2 + \delta RIR_{t-1}, \tag{C7}$$

where the notations are equal to equation (C6) except from the effect of past month's RIR ($\delta RIR_{t-1}$), where RIR is expressed in its first difference form. In line with Hammoudeh and Yuan (2008), the RIR is likely to affect the volatility of the next month and is therefore lagged by one month.

The parameter restriction of $\sum \alpha + \beta_p < 1$ must hold for all estimations, otherwise the conditional variance will be estimated using the EGARCH (1, 1) model, specified as:

$$\ln(\sigma_t^2) = \omega + \gamma_1 \frac{\varepsilon_{t-1}}{\sigma_{t-1}} + \gamma_2 \left| \frac{\varepsilon_{t-1}}{\sigma_{t-1}} \right| + \beta_1 \ln(\sigma_{t-1}^2) + \delta RIR_{t-1}, \tag{C8}$$

where the conditional variance $\sigma_t^2$ is estimated in a log-linear form to allow for positive and negative impacts, ω is the intercept, $\sigma_{t-j}^2$ is the logged past volatility effect and $\beta$ is its coefficient.

The asymmetric effect is measured by ($\gamma_1$) and the symmetric effect (replacing the ARCH effect in the original GARCH model) by ($\gamma_2$). The asymmetric effect depends on the sign of $\gamma_1$ and whether past shocks $\left(\frac{\varepsilon_{t-i}}{\sigma_{t-j}}\right)$ are positive or negative. As previously, the lag order and the Student's t-distribution have been selected based on the AIC criterion.

Estimating the parameters of the models will answer if the hypotheses defined in Section 1.2 hold for the sample, which imply testing if:

(1) the average standard deviation ($\sqrt{\sigma_t^2}$) estimated in equation (C6) is higher for the spot and the 3-months' contract relative to the 15-month's and 27-month's contract. Comparing the volatility of the contracts with the volatility of the time spread will indicate if temporary shocks and persistent shocks have been equally important to copper price volatility during the sample period.

(2) an increase in the interest rate increases the volatility of spot-returns, 3-months' returns, and the first difference of the time spread, such that $\delta > 0$ in (C7).
(3) the effect of the RIR ($\delta$) in (C6) should not be statistically significant for the 15-months' contract or the 27-months' contract.

# 5. Empirical results

## 5.1. Estimated volatilities with the GARCH (1, 1) and EGARCH (1, 1) models

Table 8 reports the empirical results of the GARCH (1, 1) model, assuming the Student's t-distribution. The first column reports the estimated parameters of each model and the associated log-likelihood and AIC values as well as the average volatility and persistence of the models. The rest of the columns display the estimates of the conditional mean and variance regressions for each variable. The impact of past shocks, the ARCH effect ($\alpha_1$), is significantly positive for all contracts implying that shocks to the log-returns increase the conditional variance in the next month. The ARCH effect on the variance of the 15-months' and 27-months' contracts are considerably larger than for the spot and 3-months' contracts. A 1% increase in past shocks to the log-return of the spot contract and 15-month's contract increase the next month's variance by 0.07% and 0.06% respectively while the ARCH effect is almost double for the 15-months' and 27-months' contracts.

The variance of all contracts is strongly dominated by past volatilities, the GARCH effect ($\beta_1$), which indicates that past month's variance is more important to the current month's variance than past shocks. The GARCH effect is significantly positive for all contracts and particularly large for the spot and 3-months' contracts for which an increase in past volatilities increases the variance of the next month by 0.89% while the effect is only 0.83% and 0.82% for the long-term contracts.

Table 8: Volatility of the log-returns of the contracts

| GARCH t-distribution: | Spot | | 3-month | | 15-month | | 27-month | |
|---|---|---|---|---|---|---|---|---|
| | Mean | Variance | Mean | Variance | Mean | Variance | Mean | Variance |
| Constant | 0.0019 | 0.0003 | 0.0020 | 0.0002 | 0.0002 | 0.0002 | -0.0002 | 0.0002 |
| | (0.0038) | (0.0003) | (0.0036) | (0.0002) | (0.0030) | (0.0002) | (0.0028) | (0.0002) |
| $Dep.var_{t-1}$ | 0.0980* | | 0.0991* | | 0.0840 | | | |
| | (0.0582) | | (0.0575) | | (0.0596) | | | |
| $\alpha_1$ | | 0.0684* | | 0.0647* | | 0.1113** | | 0.1148** |
| | | (0.0376) | | (0.0357) | | (0.0540) | | (0.0538) |
| $\beta_1$ | | 0.891*** | | 0.8920*** | | 0.8290*** | | 0.8240*** |
| | | (0.0608) | | (0.0610) | | (0.0783) | | (0.0782) |
| LL | 355.0187 | | 367.1899 | | 409.3746 | | 428.4485 | |
| AIC | -698.0374 | | -722.4 | | -806.7 | | -846.9 | |
| Volatility | 0.0721 | | 0.0697 | | 0.0605 | | 0.0574 | |
| Persistence | 0.959 | | 0.957 | | 0.9403 | | 0.9388 | |

Notes: $Dep.var_{t-1}$ denotes the dependent variable in the mean equation lagged one month, $\alpha_1$ denotes the lagged ARCH effect, and $\beta_1$ is the lagged GARCH effect. Standard errors are in parentheses. ***, **, * represent the statistical significance levels 1%, 5% and 10%, respectively. LL is the log likelihood of the model. AIC denotes the Akaike Information Criterion. The volatility is calculated as the mean of the square root of the estimated conditional variance of each variable. Persistence is calculated as the sum of the ARCH and GARCH parameters ($\alpha_1 + \beta_1$).

The persistence of the estimated variance of the models is measured by ($\alpha_1 + \beta_1$) and indicates how fast the volatility effect of past shocks and variances declines. The persistence is over 0.9 for all models, which is very high and implies that the effect of past shocks persists for a long period of time.

The estimated parameters of the time spread (Table 9) differ from the estimates of the other contracts in that the ARCH effect is much larger while the GARCH effect is in line with the long-term contracts in terms of magnitude. The ARCH and GARCH parameters sum to greater than one and the parameter restriction is thus violated. This implies that the effect of a shock will decay very slowly, or even linger on to infinity. The model is thus unstable and the EGARCH model is therefore applied to capture potential asymmetric effects of lagged positive and negative shocks on the variance of the spread. Results from the EGARCH estimation are reported in the third and fourth column in Table 9. The asymmetric effect ($\gamma_1$) is negative and considerably larger than the symmetric effect (which replaces the regular ARCH effect). A negative shock increases the volatility by 0.3935% (calculated as $\gamma_1 + \gamma_2$) in the next month while a positive shock decreases next month's volatility by -0.09%. This indicates that unanticipated negative shocks to the spread (the futures and the spot price approaches) increases the current conditional variance more than an unexpected increase in the spread in the previous month, which has a dampening effect on current volatility.

The EGARCH model was applied to the other variables as well but the coefficient estimating the asymmetric effect lacked statistical significance, implying that it is only the spread that is subject to asymmetric effects.

Table 9: Volatility of the first difference of the time spread, measured with GARCH (1,1) and EGARCH (1,1) models

| | Time Spread – GARCH (1,1) | | Time Spread – EGARCH (1,1) | |
|---|---|---|---|---|
| | Mean | Variance | Mean | Variance |
| Constant | 0.0004 | 0.0000 | -0.0000 | -0.0208 |
| | (0.0006) | (0.0000) | (0.0005) | (0.0515) |
| $\alpha_1$ | | 0.2130** | | |
| | | (0.0995) | | |
| $\beta_1$ | | 0.8246*** | | 0.9991*** |
| | | (0.0424) | | (0.0068) |
| $\gamma_1$ | | | | -0.2434*** |
| | | | | (0.0657) |
| $\gamma_2$ | | | | 0.1501** |
| | | | | (0.0697) |
| LL | 752.1411 | | 757.3512 | |
| AIC | -1494 | | -1503 | |
| Volatility | 0.0255 | | 0.0239 | |
| Persistence | 1.038 | | | |

Notes: ***, **, * represent the statistical significance levels 1%, 5% and 10%, respectively. For the EGARCH (1,1) model, if the asymmetric effect is statistically significant and negative ($\gamma_1 < 0$), it implies that a negative shock will increase the volatility by ($\gamma_1 + \gamma_2$) in the next period while a positive shock affects next period's volatility by ($-\gamma_1 + \gamma_2$).

The average volatility of each contract over the sample period is calculated as the mean value of the square root of the estimated conditional variances of each month. As can be noted in Table 8, the short-term contracts are more volatile than the long-term contracts. The volatilities are plotted in Figure 4, displaying the volatility of each month. In the beginning of the sample period, the spread is about as volatile as the other contracts but during the first part of the 2000s, the volatility decreased significantly compared to the volatility level of the other contracts. The volatility during the second half of the 2000s increased for all contacts including the spread, followed by decreasing volatilities since 2010.

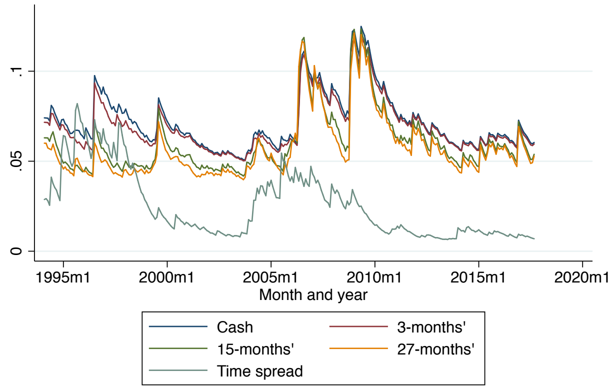


Figure 4: Estimated average volatilities during the sample period

## 5.2. Estimated volatilities including the effect of the RIR

The models including the effect of the lagged RIR in the conditional variance regression are estimated with the GARCH (1,1) specification, assuming the Student's t-distribution. Empirical results from the estimations are reported in Table 10. First, it can be noted that after including the RIR variable, the ARCH effect is no longer statistically significant. This implies that past shocks do not have an impact on current volatility. In contrast, the GARCH effect is greater than before, implying that current volatility is strongly driven by past month's volatility.

Table 10: Volatility of the log-returns, including the impact of the RIR (lagged one month)

| GARCH (1,1) t-distribution: | Spot | | 3-month | | 15-month | | 27-month | |
|---|---|---|---|---|---|---|---|---|
| | Mean | Variance | Mean | Variance | Mean | Variance | Mean | Variance |
| Constant | 0.0005<br>(0.0037) | -9.2254***<br>(0.7185) | 0.0005<br>(0.0035) | -9.2676***<br>(0.6453) | -0.0011<br>(0.0030) | -9.3619***<br>(0.5401) | -0.0013<br>(0.0028) | -9.3458***<br>(0.5515) |
| $Dep.var_{t-1}$ | 0.1015*<br>(0.0599) | | 0.1002*<br>(0.0588) | | 0.0768<br>(0.0594) | | | |
| $RIR_{t-1}$ | | 3.5111***<br>(0.5444) | | 3.5384***<br>(0.4707) | | 3.7219*** | | 3.6867*** |
| $\alpha_1$ | | 0.0268<br>(0.0271) | | 0.0123<br>(0.0230) | | 0.0197<br>(0.0309) | | 0.0343<br>(0.0368) |
| $\beta_1$ | | 0.9117***<br>(0.0382) | | 0.9232***<br>(0.0327) | | 0.8967***<br>(0.0397) | | 0.8741***<br>(0.0480) |
| LL | 361.3665 | | 374.7127 | | 418.3861 | | 435.385 | |
| AIC | -708.7 | | -735.4 | | -822.8 | | -858.8 | |
| Volatility | 0.0705 | | 0.0678 | | 0.0592 | | 0.0559 | |
| Persistence | 0.939 | | 0.936 | | 0.916 | | 0.908 | |

Notes: $RIR_{t-1}$ denotes the U.S. 3-month Treasury bill rate lagged by one month.

The effect of the RIR is significantly positively correlated with the estimated conditional variance of the variables. A 1% increase in the RIR of the previous month increases the conditional variance by more than 3.5% for all variables, which is a significant impact. The RIR effect is the largest for the 15-months' contract and the time spread (Table 11). In addition, the persistence of past variances and shocks on the conditional variance is smaller when the RIR is included in the estimation, especially for the 15-months' contract and the 27-months' contract.

The estimated parameters of the time-spread still violate the parameter restriction and the EGARCH model is therefore applied again and the results are reported in Table 11. This time, the symmetric effect (regular ARCH effect) is no longer statistically significant while the asymmetric effect is approximately of the same magnitude as before and with the same sign. This suggests that negative shocks affect the volatility of the spread but not the symmetric ARCH effect, in line with the other contracts. Interestingly, the effect of the RIR is significantly lower estimated with the EGARCH model relative to the GARCH model.

Table 11: Volatility of the first difference of the time spread, including the impact of the RIR (lagged one month)

| | Time Spread – GARCH (1,1) | | Time Spread – EGARCH | |
|---|---|---|---|---|
| | Mean | Variance | Mean | Variance |
| Constant | 0.0004<br>(0.0006) | -13.0113***<br>(1.1784) | -0.0002<br>(0.0006) | |
| $RIR_{t-1}$ | | 3.8790**<br>(1.7119) | | 0.2347**<br>(0.0949) |
| $\alpha_1$ | | 0.2201**<br>(0.1066) | | |
| $\beta_1$ | | 0.8124***<br>(0.0487) | | 0.9977***<br>(0.0056) |
| $\gamma_1$ | | | | -0.2405***<br>(0.0596) |
| $\gamma_2$ | | | | 0.0772<br>(0.0508) |
| LL | 750.1585 | | 756.8977 | |
| AIC | -1488 | | -1500 | |
| Volatility | 0.0255 | | 0.0235 | |
| Persistence | 1.033 | | | |

Notes: ***, **, * represent the statistical significance levels 1%, 5% and 10%, respectively.

Moreover, the estimated average volatility is slightly lower for all variables compared to the previous estimations except from the time spread, which has approximately the same estimated volatility. The volatilities are illustrated in Figure 5.

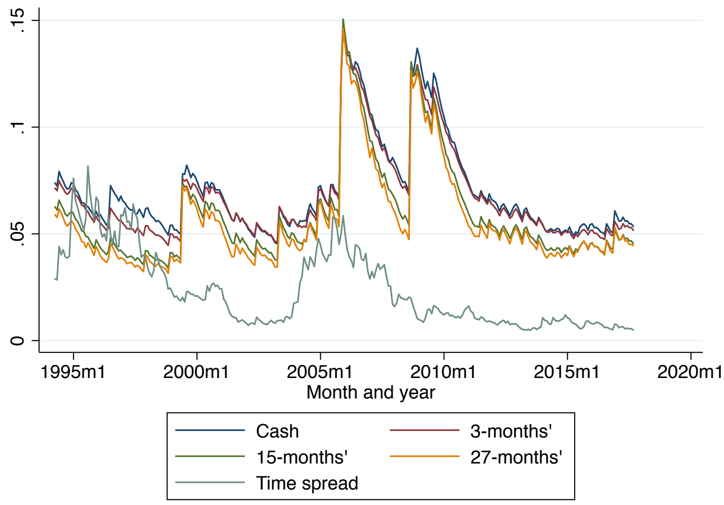



Figure 5: Estimated average volatilities during the sample period including the effect of the RIR

# 6. Discussion

The aim of this thesis was to determine if the volatility of copper prices for the past decades can be explained by the theory of storage. To test the theory, three hypotheses were outlined in Section 1.2 and specified further in Section 4.4 and the results were presented in the previous section.

Starting with the first hypothesis, which states that short-term prices are more volatile than long-term prices since they are subject to both temporary and persistent shocks, the estimated average standard deviation (volatility) of each price series were examined. In line with the theory, the spot contract is the most volatile of the contracts, followed by the 3-months' and 15-months' contracts, whereas the 27-months' contract is the least volatile. The time spread is notably less volatile than all the contracts. According to the theory, this suggests that persistent shocks have been more important to copper price volatility during the sample period, as the volatility of the time spread is only affected by temporary shocks. These results confirm the hypothesis and are also in line with the conclusion of Gruber and Vigfusson (2016), who find that the volatility of the time spread for a number of commodities is lower than the front month contract and 15-months' contract. In addition, dividing the sample into two periods and comparing the sample standard deviations of the contracts between the periods, as described in Section 4.2, shows that the volatility of all variables has increased in the second period except from the time spread which has become less volatile. The sample standard deviation of the 15-months' and 27-months' contracts had increased by 73% and 83% respectively compared to the first period, which is considerably larger relative to the short-term prices.

The overall conclusion is thus that temporary shocks have become less important while persistent shocks have become more frequent and important. Interest rates have been exceptionally low after the financial crisis in 2008, resulting in low costs of storage which should have offset short-term price fluctuations with inventory smoothing. This is demonstrated by the low volatility of the time spread while the volatility of the other contracts is higher due to being driven by persistent shocks. The fact that interest rates have been low for so long may also explain that shocks to the return has a smaller impact on the volatility in the next month compared to the volatility of the past month, which is the dominating determinant of the volatility of all contracts. In addition, the effect of shocks has a larger impact on the long-term contracts compared to the short-term, which also indicates that persistent shocks are more important than transitory ones. The finding is also in line with the conclusions of Jacks (2013) and Stürmer (2016), who identified persistent demand shocks such as growth in world GDP as the main driver of copper prices over time. The increase in persistent shocks is likely to have been driven by increased demand from emerging markets, particularly China. As the elasticity of supply of copper is low, an increase in global demand will reduce the available stocks and the consumers will bid up the price of both spot and the futures prices, causing great volatility in the market which will persist until supply responds and bring prices back to some equilibrium level.

Moreover, unexpected declines in the time spread increases the volatility in the next month and the effect is greater compared to an unexpected increase in the spread. This could be explained by the fact that a decreasing spread is often associated with increased uncertainty of future stocks of the commodity. Alquist and Kilian (2010) showed that increased uncertainty of future supply shortages causes the spread between oil futures and spot prices to decline. Increased uncertainty also causes precautionary demand for oil to increase which results in an immediate increase in the real spot price. Their conclusion is that a decrease in the spread can be interpreted as an indicator of volatility in the spot price, driven by shocks attributable to precautionary demand. The same conclusion is likely to be true for the futures—spot spread of copper prices, that is, a fall in the spread is associated with increased uncertainty of future stocks, which increases short-term volatility due to a spiking precautionary demand. This results in an increase in the volatility of prices in the next month and the effect is larger than in response to an increase in the spread, which actually decreases volatility in the next month as a rise in the spread is not associated with any uncertainty for future supply.

Furthermore, the second hypothesis states that a negative change in the RIR should decrease the volatility of short-term prices and the time spread, as low interest rates is synonymous with low storage costs and thus allows for inventory smoothing over short-term price fluctuations. The hypothesis is accepted as the RIR has a significantly positive effect on the volatility of the spot and the 3-months' contracts as well as the time spread. When the RIR is included in the estimation, the effect of past shocks to the volatility in the next month is no longer statistically significant, implying that copper price volatility is only driven by past month's volatility and changes in the RIR. The RIR has been about zero since 2009 and this could eliminate the volatility effect of a past shocks since one can build inventories at essentially no cost and shocks causing short-term fluctuations in prices can be dampened by consuming or selling out of inventories. When accounting for asymmetry in the effect of past shocks on the volatility of the time spread, the estimated effect of the RIR is lower than when estimated with the GARCH model. The explanation to this is not obvious but the dampening effect of the RIR could be conditional on the sign of the shock.

Finally, the third hypothesis, must be rejected as the effect of the RIR is statistically significant for all contracts, including the long-term prices. According to the theory, lower interest rates should only decrease the volatility of the spread and the short-term prices as those are subject to temporary shocks. The result is in contrast to the finding of Gruber and Vigfusson (2016), who conclude that the pattern of statistical significance in their empirical results supports the theory. Their results demonstrate a statistical significant effect of the RIR on the time spread of some commodities while the effect on the short- and long-term prices is not statistically significant. The empirical results of this thesis show the opposite: the interest rate decreases the volatility of not only the time spread and the short-term prices but the long-term prices as well. Again, this could be due to the extremely low level of the RIR for a long period of time and market participants may thus have expected the situation to persist, resulting in high levels of inventories and hence the dampening effect of the RIR even on long-term price volatility. Estimating volatility with the GARCH model is however limited to explaining the average effect of the RIR over the sample period and analysing if the effect of the RIR changes over different time horizons of the sample is unfortunately not possible.

Another weakness of the study is that, as described by Akram (2008), shocks that increases the future price of a commodity, such as higher economic growth, could also result in higher real interest rates. It could thus be that the volatility of long-term copper prices increases with increasing interest rates due to being correlated with growth in world GDP. An increase in the RIR would then increase the volatility of long-term prices as it constitutes a demand shock expected to persist. A single equation approach such as the GARCH and EGARCH model cannot account for such a dynamic relationship between interest rates and copper prices. One approach would be to apply the multivariate GARCH model, which allows the current volatility of a time series variable to be influenced not only by its own past errors and volatility but also by past values of other time series. This extended analysis is beyond the scope of this thesis but could be a methodology for future research on why the effect of the RIR differ between studies and maturity of the contract.

Additionally, the thesis is limited to examining the effect of the RIR. There may be other omitted variables that are important to the volatility of copper prices. Moreover, using prices quoted on the LME and the U.S. interest rate may not be valid for all copper markets. Even though copper futures traded on the LME are said to represent global prices, empirical results using copper futures prices traded on the Shanghai Futures Exchange could differ from the results of this study. Examining if the same pattern is true when using other data series could also be an interesting topic for future research.

Finally, the interpretation of the empirical results relies on the assumption that the RIR affects the volatility of the variable solely through encouraging inventory smoothing, since that is the prediction of the theory of storage. However, the study does not examine whether actual levels of inventory on the market increase in response to a decrease in the RIR and hence the dampening effect of the price volatility of commodities. But as explained above, this dampening effect could be related to the RIR being correlated with world GDP and the effect may thus not be entirely associated with the predictions of the theory.

# 7. Summary and concluding remarks

The thesis has examined to what extent the predictions derived from the theory of storage can explain the volatility of copper prices during the period of 1994 to mid-2017. The theory of storage is well established in the literature, explaining how decreasing interest rates increases commodity prices by lowering the cost of storage and increasing the demand for inventories. The theory also suggests that decreasing interest rates should reduce the volatility of commodity prices, since lower storage costs contributes to higher levels of inventories which can be used to offset price fluctuations caused by temporary shocks to supply or demand. In contrast, interest rates should have no effect on long-term prices as inventory smoothing is only profitable over a short period of time. Previous research has proven these predictions to be valid for several commodities but not for copper prices. The focus of this thesis was therefore to examine this further by applying the predictions to copper price volatility exclusively.

To test if the theory holds for copper prices, temporary shocks were identified through movements in the spread between the 15-months’ futures contract and the spot price for copper. The effect of persistent shocks is represented by movements in long-term prices in terms of the 15-months’ and 27-months’ contracts. The empirical results suggest that most of the characteristics of copper price volatility can be explained by the theory: short-term prices are more volatile than long-term prices and the real interest rate has a significantly dampening effect on price volatility. In addition, the theory allows for the interpretation that the volatility of copper has been largely driven by persistent shocks in recent years. This is in contrast to previous research which have found interest rates not to affect the volatility of copper prices (Hammoudeh and Yuan, 2008; Gruber and Vigfusson, 2016). The empirical findings also show that the interest rate decreases the volatility of long-term prices, which is not consistent with the theory. The long period of exceptionally low-interest rates may explain why the results differ from the predictions of the theory. A suggestion for future research is to account for the possible correlation between interest rates and growth in world GDP in order to determine if the interest rate effect is only attributable to the cost of storage.

The overall conclusion is that the volatility of copper prices to a large extent can be explained by the theory but the reason why interest rate changes affect prices at all horizons remains unexplained. The results suggest that if the expectations of growing demand for copper and increasing interest rates are realized in the years ahead, the volatility of copper is likely to increase considerably. This will have important implications to a number of countries and industries, such as the growing sector of renewable energy systems and technologies, which rely heavily on copper.

# 8. References


Algieri, B. (2014) “The influence of biofuels, economic and financial factors on daily returns of commodity futures prices”, Energy Policy 69, pp. 227-247

Alquist, R. and E. Arbatli. (2010): Crude Oil Futures: A Crystal Ball?, Bank of Canada Review, Spring 2010. pp. 3-11

Alquist, R. and Kilian, L. (2010): What do we learn from the price of crude oil futures:, Journal of Applied Econometrics, 25, pp. 539-573

Andrews, D. W. K. (1993) “Tests for parameter instability and structural change with unknown change point”, Econometrica 61, pp. 821–856

Arbatli, E. (2008) “Futures markets, oil prices and the intertemporal approach to the current account”, Bank of Canada Working Paper No. 2008-48

Barsky, R. and Kilian, L. (2002) “Do we really know that oil caused the great stagflation? A monetary alternative, NBER Macroeconomics Annual (16), pp. 137–198

Bollerslev, T. (1986) “Generalized Autoregressive Conditional Heteroskedasticity, Journal of Econometrics, 31, pp. 307-327

Box. G.E,P., Jenkins, G.M., and Reinsel, G.C. (2008) ”Time series analysis: Forecasting and control”, 4th edition, Wiley, Hoboken, NJ

Brennan, M.J. (1958) The Supply of Storage. The American Economic Review 48 (1), pp. 50-72

Brunetti, C. and Gilbert, C.L. (1995) “Metals price volatility, 1972–1995”, Resources Policy 21(4), pp. 237-254

Cafiero, C., Bobenrieth, E., Bobenrieth, J. and Wright, BD. (2011) “The Empirical Relevance of the Competitive Storage Model”, Journal of Econometrics, 162 (1), pp. 44–54

Carter, C., Rausser, G. and Smith, A. (2011) “Commodity booms and busts”, Annual Review of Resource Economics 3, pp. 87–118

Chernenko S.V., Schwarz K.B. and Wright J.H. (2004) “The information content of forward and futures prices”, International Finance Discussion Paper No. 808, Federal Reserve Board

Dueker, M. J. (1997): “Markov switching in GARCH processes and mean-reverting stock-market Volatility”, Journal of Business & Economic Statistics, 15 pp. 26-34

Dickey, D. A. and Fuller W. A. (1979) “Distribution of the estimators for autoregressive time series with a unit root”, Journal of the American Statistical Association (74), pp. 427–431

Engle, R. (1982) “Autoregressive conditional heteroskedasticity with estimates of the variance of the U.K. inflation, Ekonometrica 50, pp. 987-1008

Engle, R. (2001) “GARCH 101: The use of ARCH/GARCH models in applied econometrics”, Journal of Economic perspectives, Vol.15 (No.4), pp. 157-168

Erb, C.B. and Campbell H.R. (2005), "The Tactical and Strategic Value of Commodity Futures", Financial Analysts Journal, vol. 62, no. 2, pp. 69–97

Fama, E. and French, K. (1987) “Commodity futures prices: some evidence on forecast power, premiums, and the theory of storage”, The Journal on Business, Vol. 60 (1), pp. 55–73

Fama, E. and K. French (1988) “Business Cycles and the Behavior of Metals Prices”, Journal of Finance (43), pp. 1075-1093

Frankel, J. (2008) “The Effect of Monetary Policy on Real Commodity Prices”, Asset Prices and Monetary Policy, John Y. Campbell, University of Chicago Press, pp. 291-333

Frankel, J. (2014) "Effects of speculation and interest rates in a “carry trade” model of commodity prices," Journal of International Money and Finance, 42(C), pp. 88-112

Frankel, J.A. and Hardouvelis, G.A. (1985) “Commodity Prices, Money Surprises and Fed Credibility”, Journal of Money, Credit and Banking, 17 (4), pp. 425-438

Geman, H. and Ohana, S. (2009) “Forward Curves, Scarcity and Price Volatility in Oil and Natural Gas Markets”, Energy Economics (31), pp. 576-585

Geman, H. and Smith, W.O. (2013) “Theory of storage, inventory and volatility in the LME base metals”, Resources Policy, 38, pp. 18-28

Gorton, G and Rouwenhorst, G.K. (2006) “Facts and Fantasies about Commodity Futures.” Financial Analysts Journal, vol. 62, no. 2, pp. 47–68

Granger, C.W.J., and Newbold, P (1974) “Spurious regressions in econometrics”, Journal of Econometrics (2), pp. 111-120

Greer, R. (2000) “The Nature of Commodity Index Returns”, Journal of Alternative Investments, vol. 3, no. 1, pp. 45–52.

Gruber, J. and Vigfusson, R. (2016) “Interest Rates and the Volatility and Correlation of Commodity Prices”, Macroeconomic Dynamics, pp. 1-20

Hansen, P. and Lunde, A. (2004) ”A Forecast Comparison of Volatility Models: Does Anything Beat a GARCH(1,1) Model?”, Journal of Applied Econometrics 20, pp. 873–889.

Hamilton, J. (2008) “The new Palgrave dictionary of economics”, chapter Oil and the Macroeconomy, MacMillan, London

Hammoudeh, S. and Yuan.Y. (2008) “Metal volatility in presence of oil and interest rate shocks”, Energy Economics (30), pp. 606–620

International Monetary Fund (2014) “World Economic Outlook: Legacies, Clouds and Uncertainties”, pp. 25

Jacks, D. S. (2013) “From Boom to Bust: A Typology of Real Commodity Prices in the Long Run” NBER Working Paper No. 18874 (Cambridge, Massachusetts: National Bureau of Economic Research)

Kaldor, N. (1939) “Speculation and Economic Theory,” Review of Economic Studies (7), pp. 1-27

Karali, B. and Power, G.J (2013) “Short- and long-run determinants of commodity price volatility, American Journal of Agricultural Economics, 95 (2), pp. 1-15

Keynes, J.M. (1930) “A treatise on money”, Harcourt, Brace and Company, New York

Kornher, L. and Kalkuhl, M. (2013) “Food price volatility in developing countries and its determinants”, Quarterly Journal of International Agriculture, 52, pp. 277-308

Ma, J., Nelson, C.R., and Startz, R. (2006) ”Spurious inference in the GARCH (1, 1) model when it is weakly identified, Studies in nonlinear dynamics and econometrics, 11(1), Article 1

Maddala, G.S. and In-Moo Kim (1998) “Unit roots, cointegration and structural change”, Cambridge University Press

Mandelbrot, B. B. (1963) ”The variation of certain speculative prices”, Journal of Business 36, pp. 394–419.

Meucci, A. (2005) “Risk and Asset Allocation”, Springer-Verlag Berlin Heidelberg

Miller, R. G. (1997) ”Beyond ANOVA: Basics of Applied Statistics”, London: Chapman & Hall

Nelson, D. B. (1991) ”Conditional heteroskedasticity in asset returns: A new approach”, Econometrica 59 (2)2004) , pp. 347–370

Ng, V. and Pirrong, S. (1994) “Fundamentals and Volatility: Storage, Spreads, and the Dynamics of Metals Prices”, Journal of Business (67), pp. 203-230

Schwartz, E.S. and Smith J.E. (2000) “Short-term variations and long-term dynamics in commodity prices”, Management Science, 46 (7), pp. 893-911

Stürmer, M. (2016) “150 Years of Boom and Bust - What Drives Mineral Commodity Prices”, Macroeconomic Dynamics, pp. 1-16

Svedberg, P. and Tilton, J. (2006) “The real, real price of nonrenewable resources: copper 1870-2000”, World Development, 34(3), pp. 501–519

Svensson, LEO. (2005) “Oil prices and ECB monetary policy”, Mimeo, Department of Economics, Princeton University

Stock. J.H. and Watson M.M. (2012) “Introduction to econometrics”, Third Edition

Tang, K. and Xiong, W. (2012) “Index Investment and Financialization of Commodities,” Financial Analysts Journal, 68 (6)

Working, H. (1933) “Price Relations Between July and September Wheat Futures at Chicago Since 1885”, Wheat studies of the Food Research Institute, 9(6)

Working, H. (1934), “Price Relations Between May and New-Crop Wheat Futures at Chicago Since 1885. Wheat Studies of the Food Research Institute, 10 (5)

Working, H. (1948) “Theory of the Inverse Carrying Charge in Futures Markets”, Journal of Farm Economics, 30 (1), pp. 1-28.

Working, H. (1949) “The theory of Price of Storage”, The American economic review, Vol. 39 (6), pp. 1254-1262

Zivot, E. (2009) “Practical issues in the analysis of univariate GARCH models”, Handbook of Financial Time Series pp. 113-155

# 9. Appendices

## 9.1. Appendix 1: Histograms of the distribution of the variables compared to the normal distribution

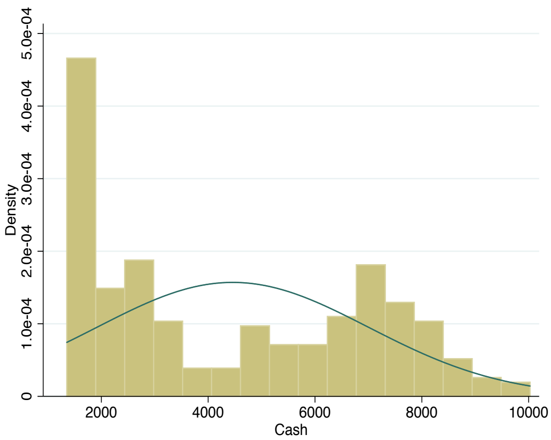



Figure 1. a) Distribution of the log cash price

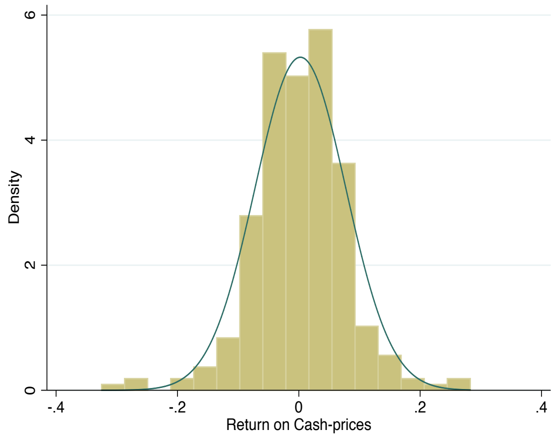



Figure 1. b) Distribution of the log-return cash price

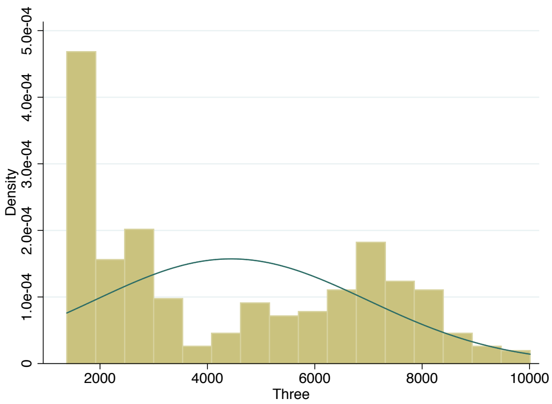



Figure 1. c) Distribution of the log 3-months' price

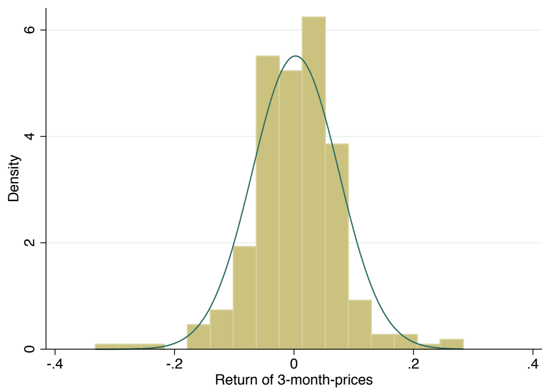



Figure 1. d) Distribution of the log-return 3-months' price

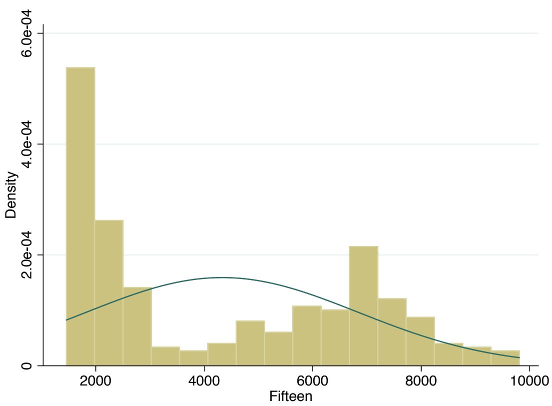



Figure 1. e) Distribution of the log 15-months' price

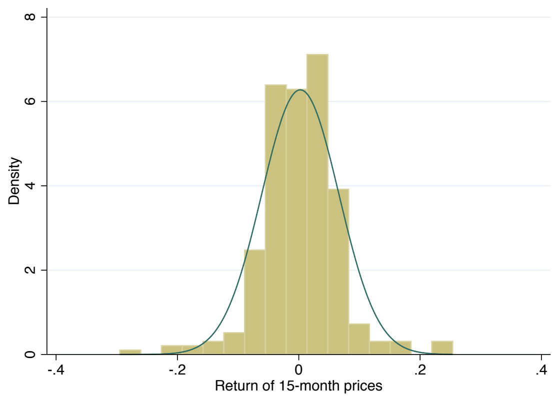



Figure 1. f) Distribution of the log-return 15-months' price

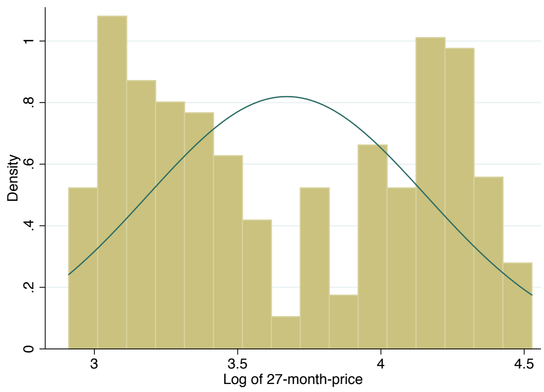



Figure 1. g) Distribution of the log 27-months' price

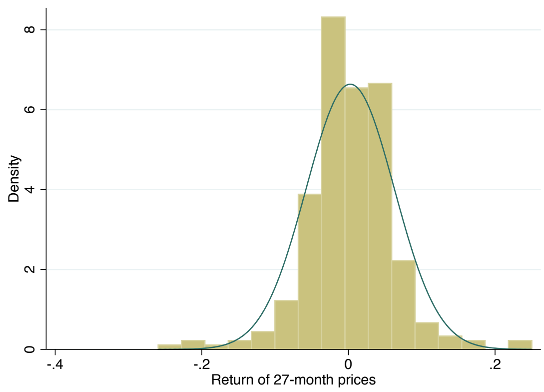



Figure 1. h) Distribution of the log-return 27-months' price

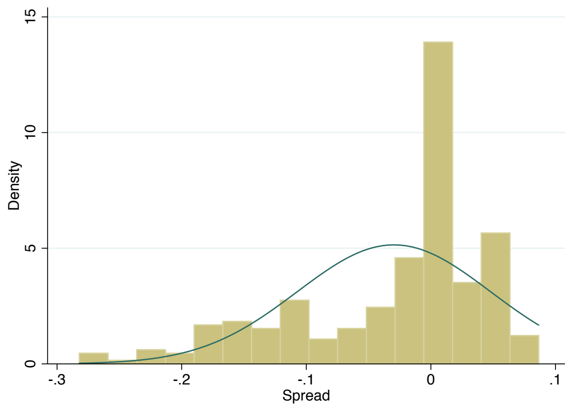



Figure 1. i) Distribution of the time spread

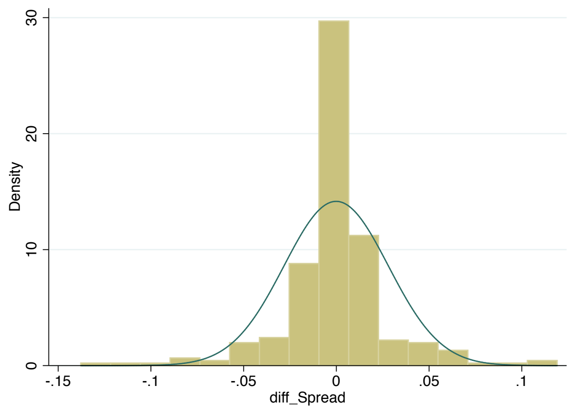



Figure 1. j) Distribution of the first differenced time spread

## 9.2. Appendix 2: Time series plots of the prices

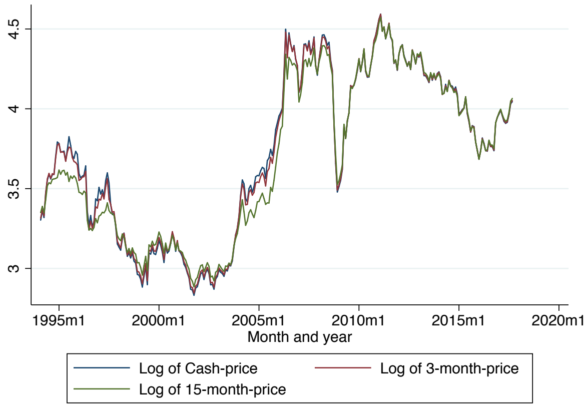



Figure 2.a) Log-prices of the contracts

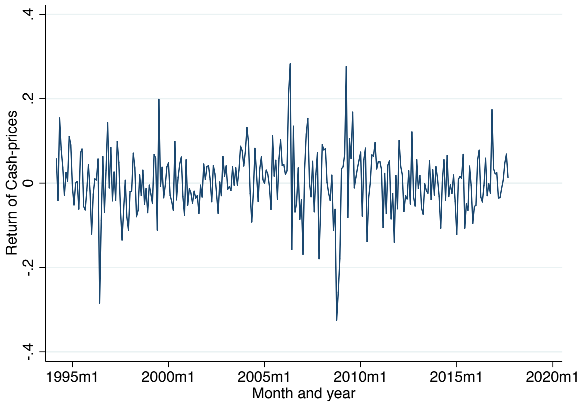



Figure 2.b) Log-returns of the cash contract

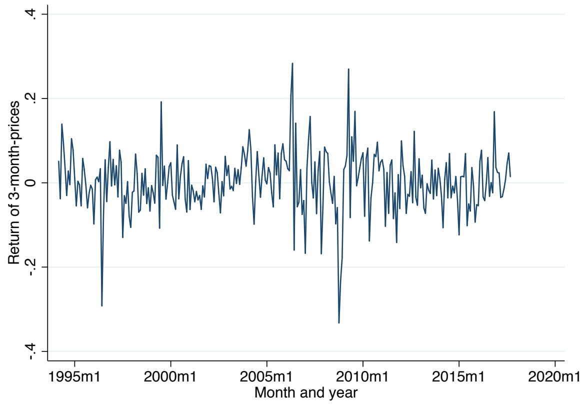



Figure 2.c) Log-returns of the 3-months' contract price

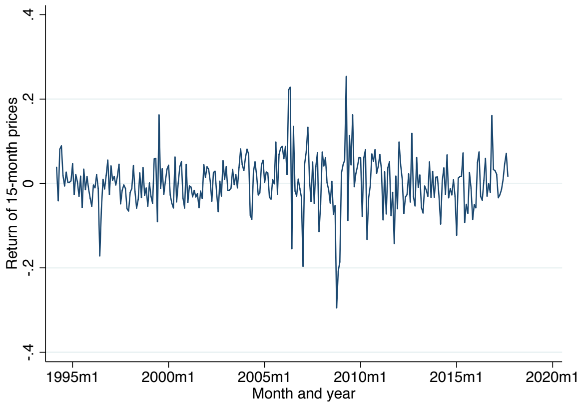



Figure 2.d) Log- returns of the 15-months' contract

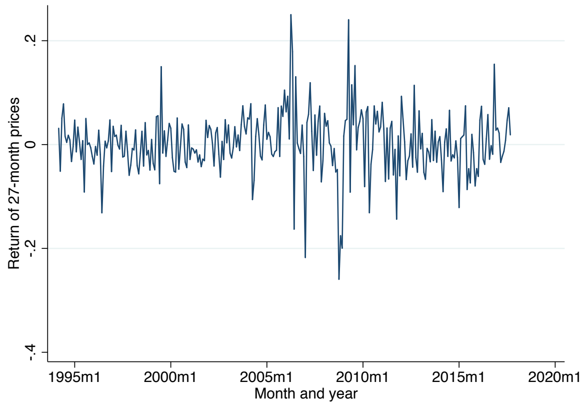



Figure 2.e) Log-return of the 27-months contract

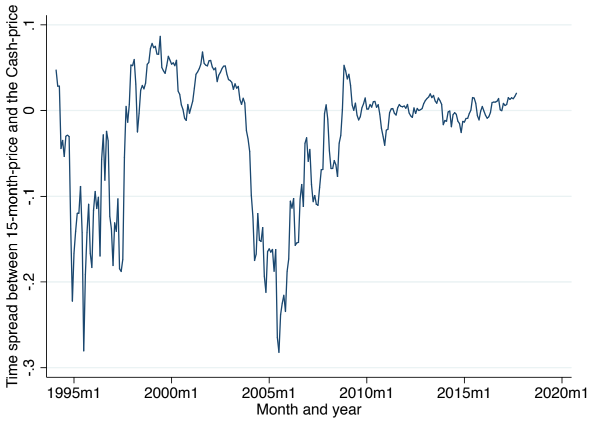



Figure 2.f) Time spread

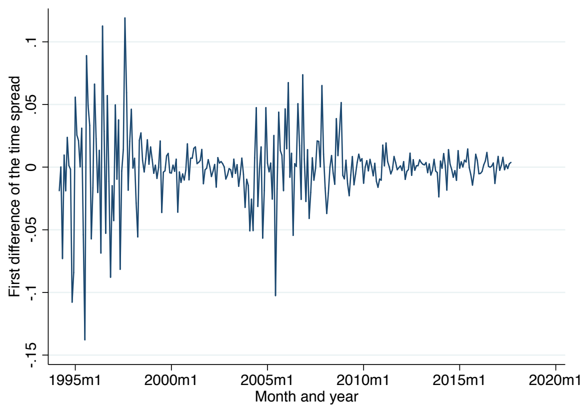



Figure 2.g) First difference of the time spread

## 9.3. Appendix 3: Time series plots of the residuals of the mean models

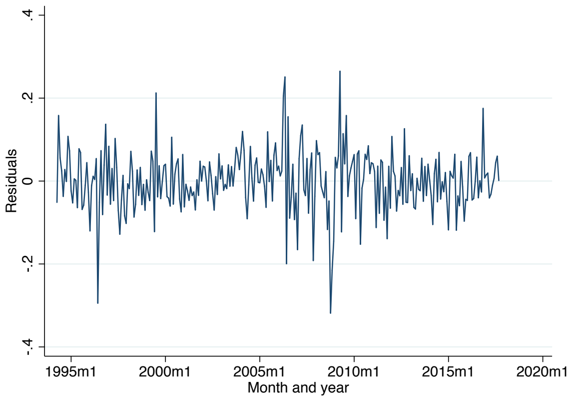



Figure 3. a) Residuals of the mean cash model

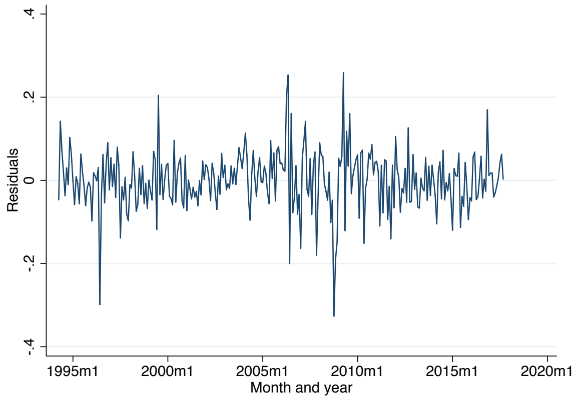



Figure 3. b) Residuals of the mean 3-months' model

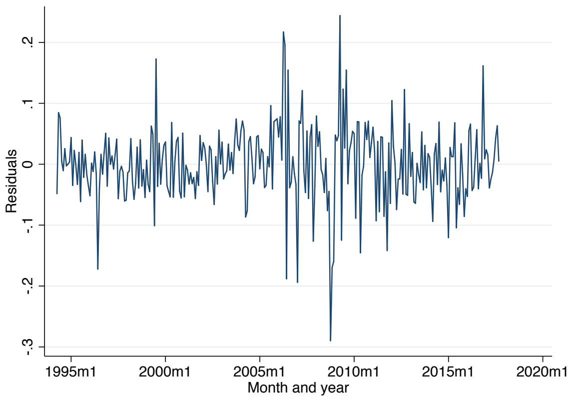



Figure 3. c) Residuals of the mean 15-months' model

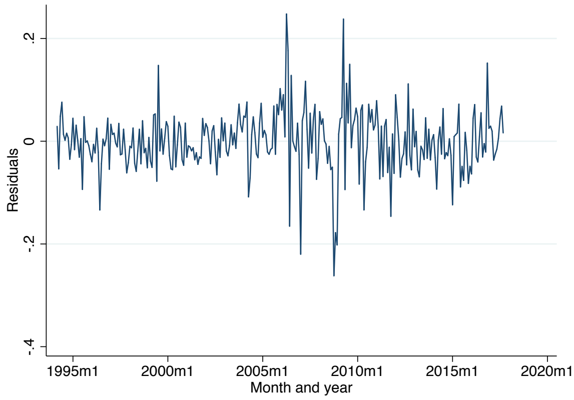



Figure 3. d) Residuals of the mean 27-months' model

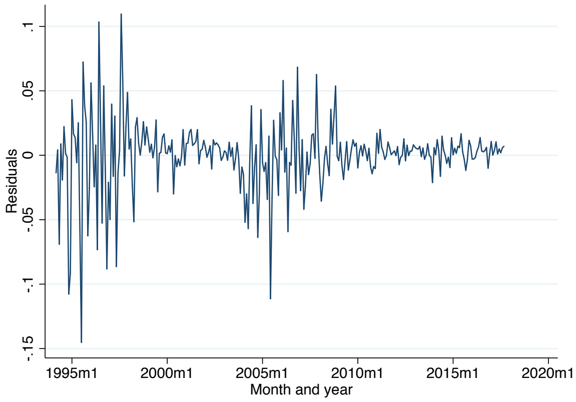



Figure 3. e) Residuals of the mean time spread model

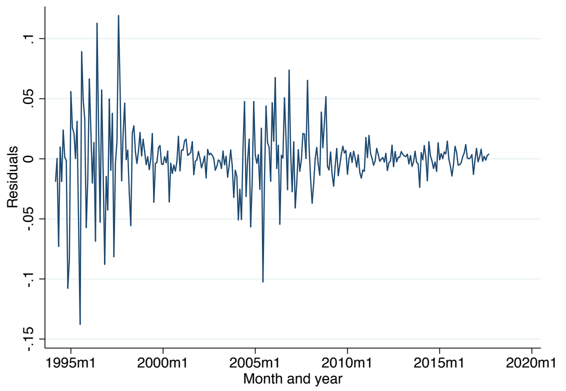



Figure 3. f) Residuals of the mean first difference time spread model

## 9.4. Appendix 4: Autocorrelation (AC) and partial autocorrelation (PAC) of the log-returns and their lags

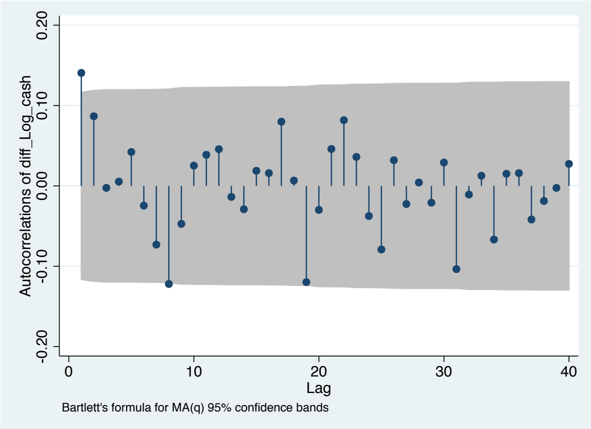



Figure 4. a) Autocorrelation (AC) of the cash-price

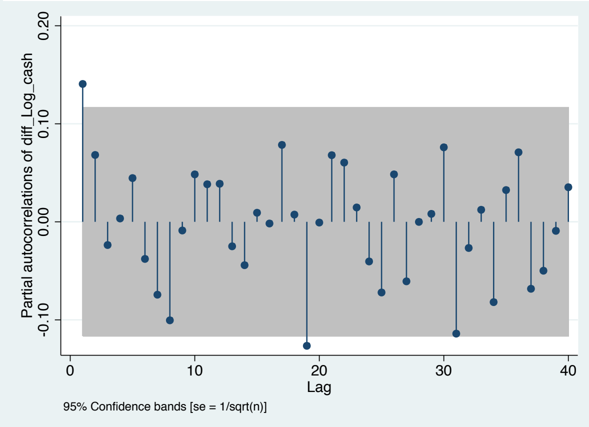



Figure 4. b) Partial autocorrelation(PAC) of the cash-price

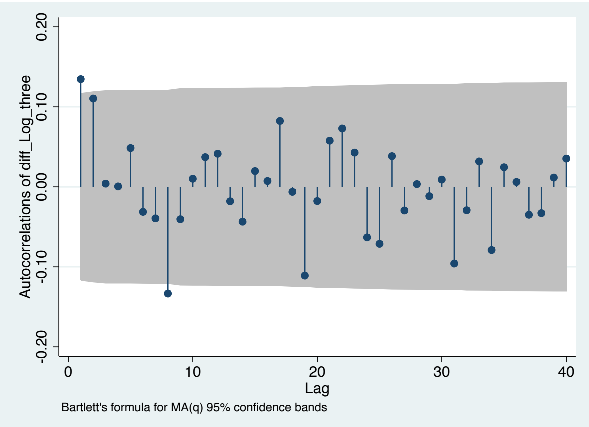



Figure 4. c) AC of the 3-months' contract

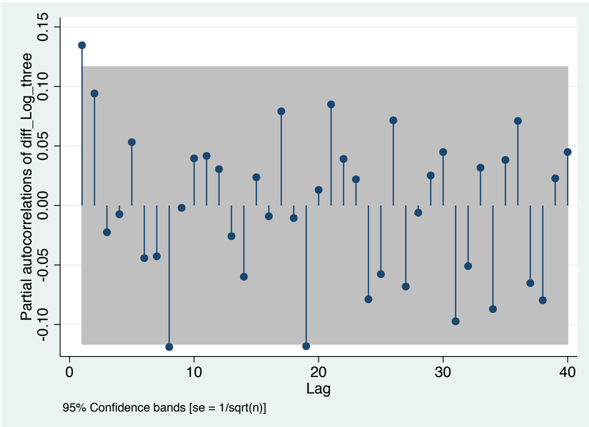



Figure 4. d) PAC of the 3-months' contract

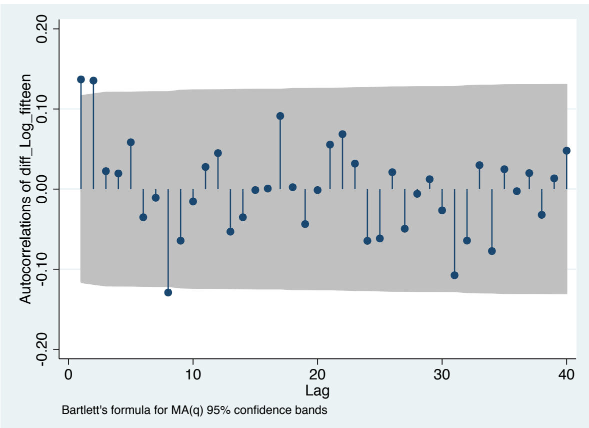



Figure 4. e) AC of the 15-months' contract

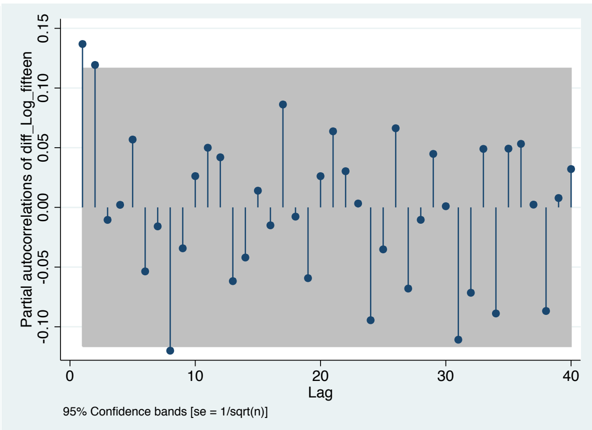



Figure 4.f) PAC of the 15-months' contract

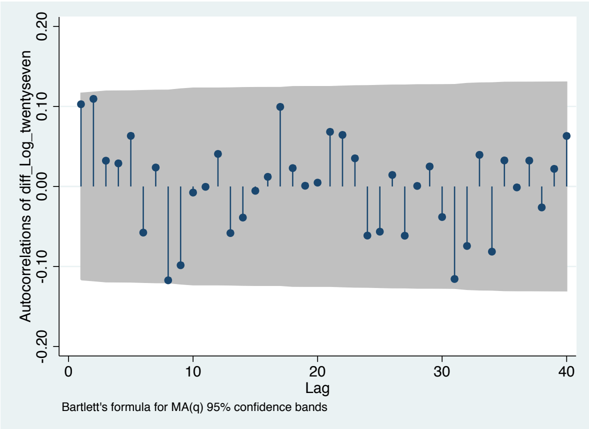



Figure 4. g) AC of the 27-months' contract

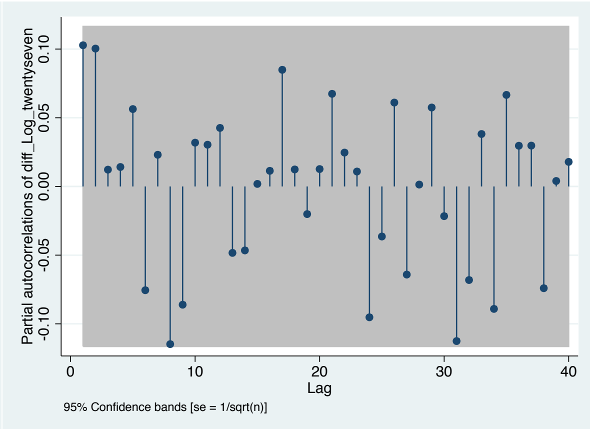



Figure 4. h) PAC of the 27-months' contract

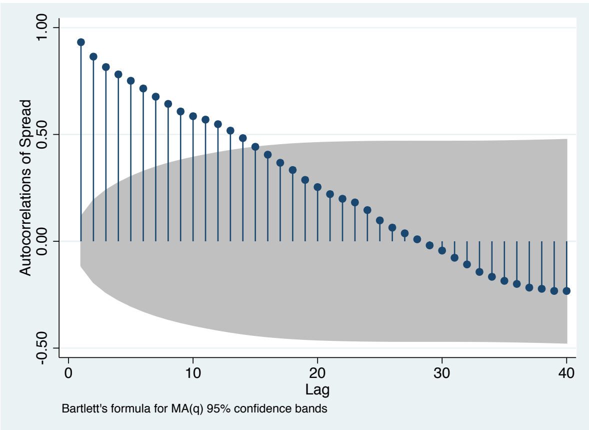



Figure 4. i) AC of the time spread

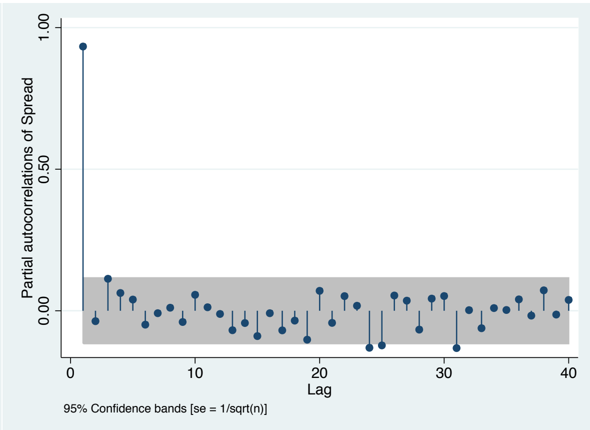



Figure 4. j) PAC of the time spread

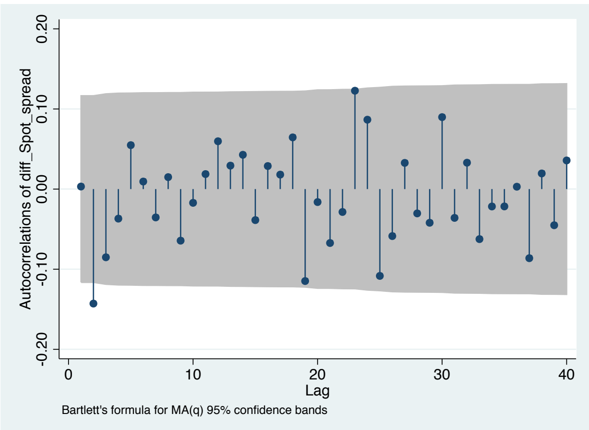



Figure 4. k) AC of the first difference of the time spread

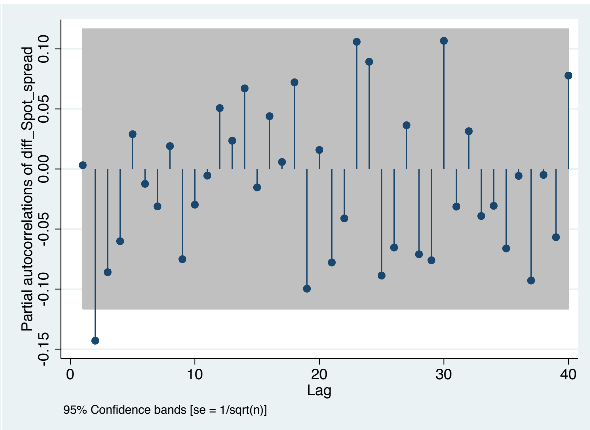



Figure 4. l) PAC of the first difference of the time spread

### 9.5. Appendix 5: Standardized normal probability plots (P-P plots) and quantile-normal plots (Q-Q plots) of the residuals of the mean model

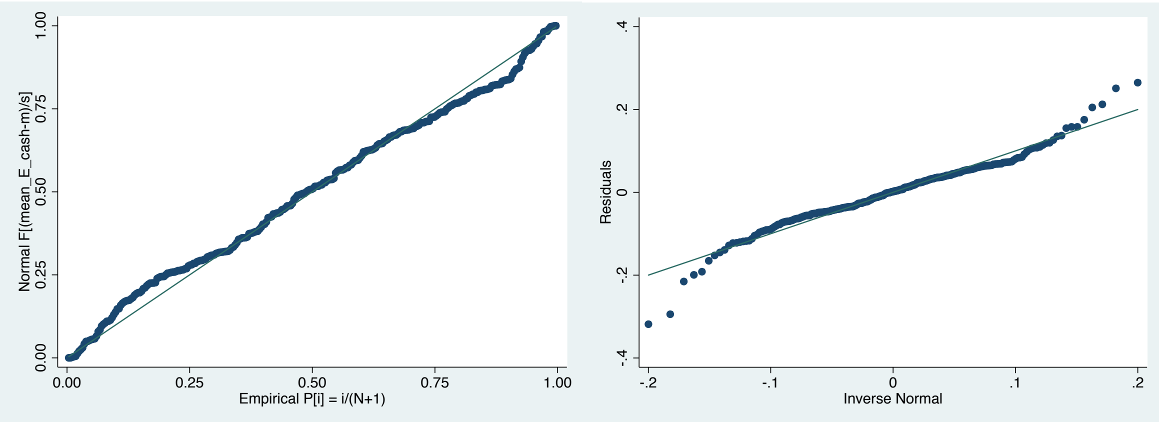



Figure 5. a) P-P plot and Q-Q-plot of the residuals of the mean cash model

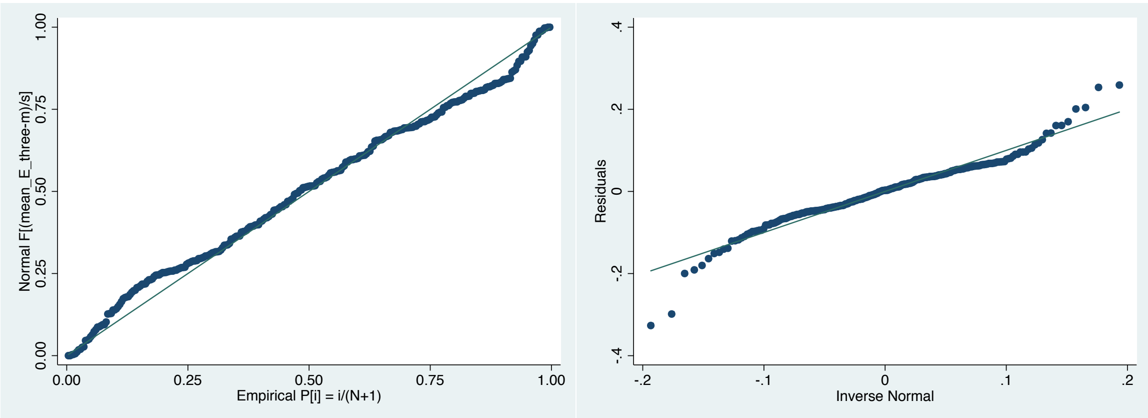



Figure 5. b) P-P plot and Q-Q-plot of the residuals of the mean 3-months' model

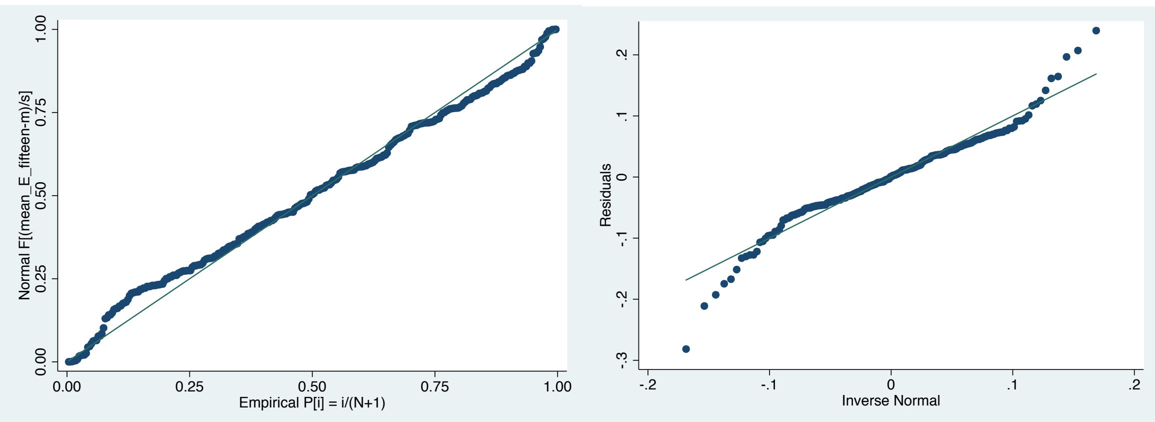



Figure 5. c) P-P plot and Q-Q-plot of the residuals of the mean 15-months' model

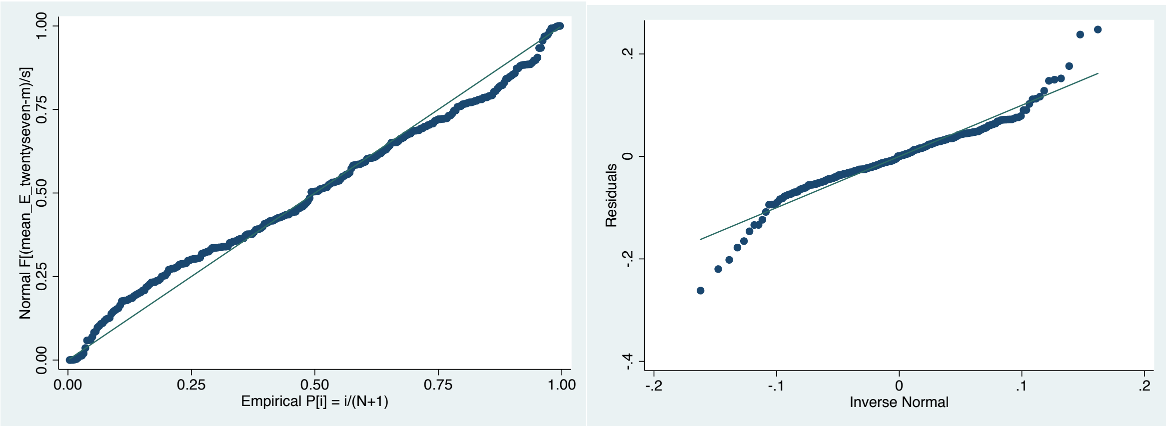



Figure 5. d) P-P plot and Q-Q-plot of the residuals of the mean 27-months' model

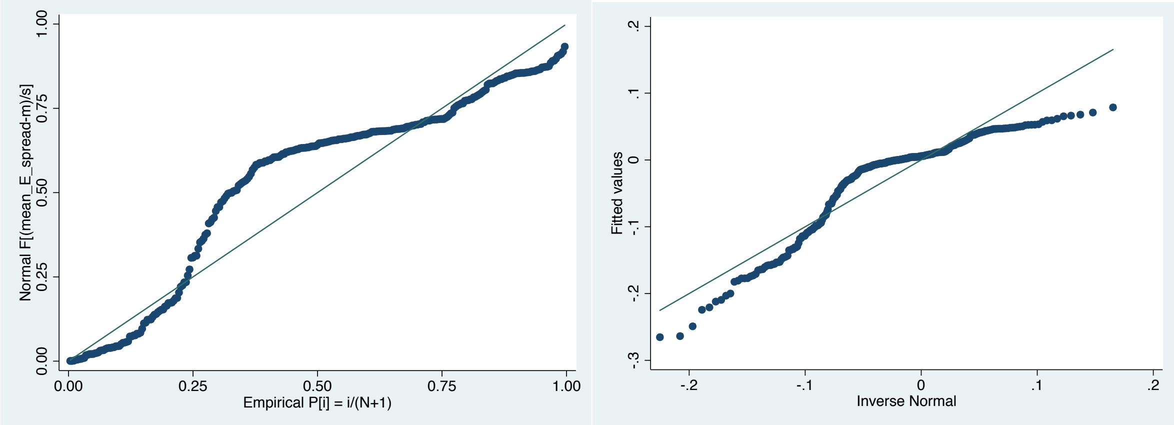



Figure 5.

e) P-P plot and Q-Q plot of the residuals of the mean time spread model

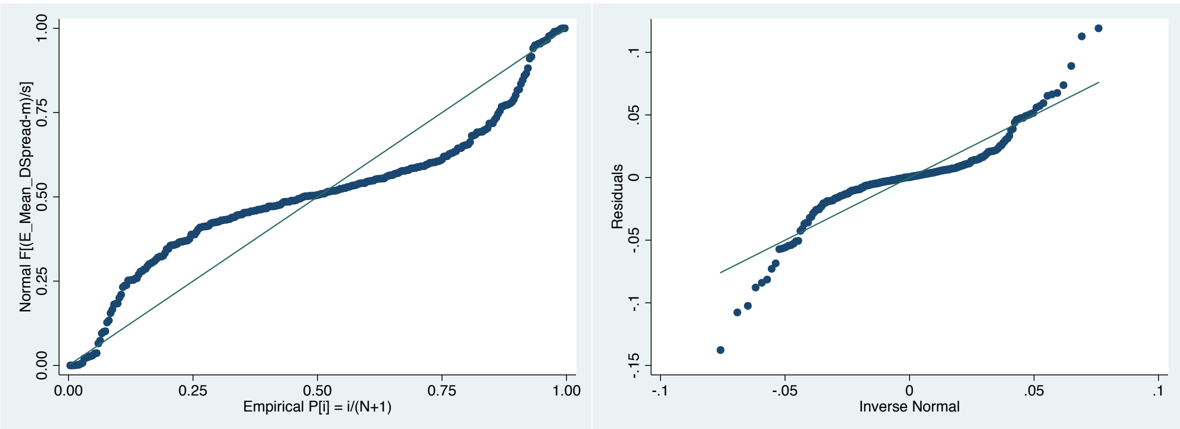



Figure 5. f) P-P plot and Q-Q plot of the residuals of the mean first difference of the time spread model