Impact of Financialization on the Roll Return of Commodities

Master Thesis

Tilburg University, School of Economics and Management

Finance-CFA Track

Artem Vdovenko

ANR: 262202

Supervisor: F. A. De Roon

Publication Date: August 2013

Table of Contents

Abstract

Despite a rapid growth of commodity spot prices in recent times, the total commodity return of passive investors is substantially lower than one could have expected due to the negative roll return. Many practitioners blame the activity of index traders and speculators for the negative impact on the roll return. This paper investigates the impact of both these trader groups during the period when the most popular indices, the GSCI and the DJ-UBSCI, roll their positions. The autoregressive distributed lag model is applied to examine the impact of non-stationary level variables. It is found that index traders decrease the roll return during the rolling period. However, it is also observed that speculators do not influence the roll return during the rolling period. Instead, they push the spread of futures down before the rolling period and provide the market with the liquidity during the rolling period. These results imply that, in order to minimize the negative roll return, index traders should seek for the indices investing in the most liquid commodity futures and rolling at the most liquid time. Such strategy allows them to decrease the share of index traders that is rolled periodically and mitigate its negative impact.

Introduction

A new type of investor, passive investors have increased rapidly their inflows in the commodity market. In the five-year period from 2003 to 2008, the assets under management of index investment funds rose from $ 15 B to $ 200 B (Rossi, 2011). Furthermore, the assets continued increasing from $ 200 B to $ 400B during the next 5 years[1]. Referring to these changes, Domanski and Heath introduce a new term that reflects the behavior of the commodity market after 2004; they refer to this as financialization of the commodity market (Domanski & Heath, 2007). Initially, the commodity market was used by producers and consumers for hedging against volatility of spot prices. The second type of market participants, speculators, provided the market with liquidity. The impact of index traders on prices and term structure is under a heavy debate among the academic community. Some researchers claim that index traders are responsible for a boom in commodity prices in 2008. Others argue that a rise of the commodity prices is accounted for by a sharp increase of demand and inelastic supply in the physical market[2].

The level of the futures price is not of a primary interest of the institutional investors. This paper considers the commodity markets in a view of institutional and, more generally speaking, passive investors. The futures return and the roll return, which is determined by the term structure, attract investors. The negative spread[3] has a detrimental effect on returns of both institutional and individual investors that need to roll contracts. In 2012, prices of natural gas futures had 24% growth. Instead, passive investors lost one-fifth of their investment in natural gas[4] over 2012. Figure 1 illustrates the impact of the roll returns during the period from 1999 to 2013. The price of the nearest futures contracts of the components of the GSCI has increased after 2009, which is reflected by the graph of the GSCI spot index. At the same time, the graph of total returns is almost flat during the period from 2009 to 2013. According to Mou (2011), the negative effect of rolling exceeds total management fees of funds over the period from 2000 to 2009.

While most of the literature regards the impact of financialization on commodity prices, this paper focuses on the impact of financialization on the roll return or the spread of futures prices. There a few articles related to this question. Brunetti and Reiffen show a negative impact

---

[1] The commodity research of Barclays Capital for the 19th of April of 2013, "The Commodity Refiner From an age of shortage to an era of enough" http://www.stockmarketnews.biz/wp-content/uploads/2013/04/Commodities-Research9.pdf

[2] Irwin and Sanders published an elaborate review of the most important articles about the financialization impact on the commodity market (Irwin & Sanders, Index Funds, Financialization, and Commodity Futures Markets, 2011).

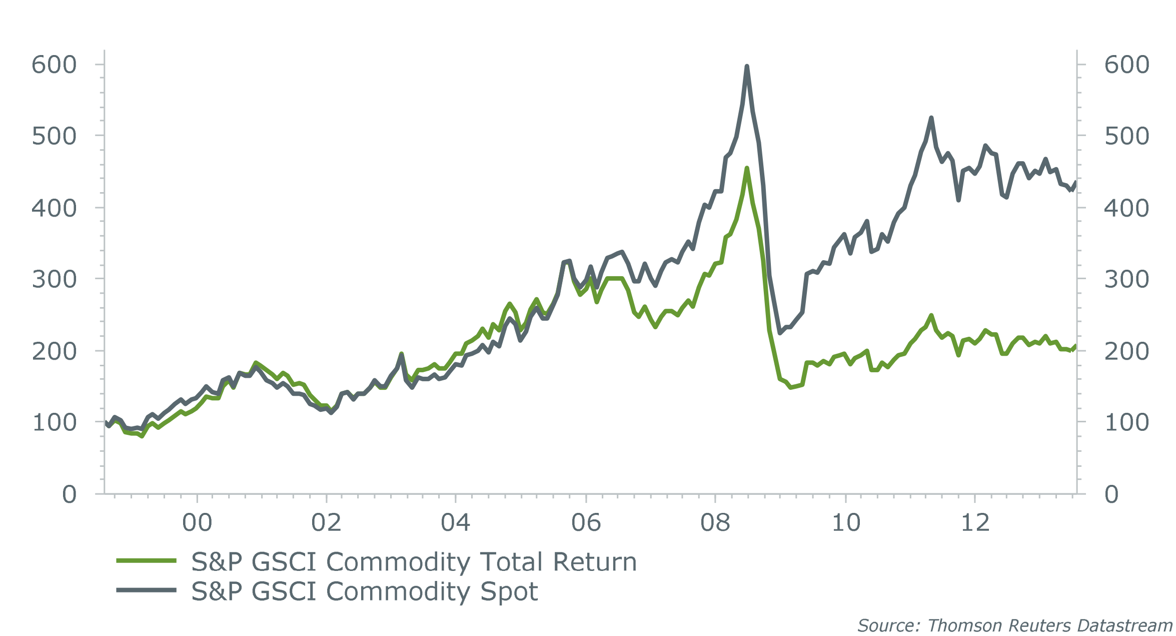
[3] Spread of futures contracts as a price difference between the nearest-to-maturity contract and next nearest-to-maturity contract and roll yield (return) are used interchangeably in this article.

[4] http://online.wsj.com/article/SB10001424127887324073504578109402867864658.html

of the share of index traders to open interest (Brunetti & Reiffen, 2011). Index traders push the price of a nearest-to-maturity contract down and push the price of a next nearest-to-maturity contract up. The effect is that the roll return goes down. However, this paper analyses only three agricultural commodities over the period from 2003 to 2008.

Figure1. The graph of the GSCI total return and the GSCI spot component, 1999-2013, Datastream

The total return index measures the returns accrued from investing in fully-collateralized nearby commodity futures, and the spot index measures the level of nearby commodity prices.



Several articles observe a decrease of the spread during the rolling period. Index investors invest mainly in the GSCI and DJUBSCI, which are rolled almost in the same period, which is called as the Goldman roll period in the thesis. Mou (Mou, 2011) investigates the profitability of the front-running strategy that takes a short position of a nearest-to-maturity contract and a long position of a next contract before the Goldman roll and close these positions after the Goldman roll. The author observes that the strategy has a high Sharpe ratio and positive significant returns. This result indicates that index traders influence the spread during the rolling period. In addition, it also implies that speculators may take arbitrage of the activity of index traders and earn profits. But the impact of speculators has not been investigated properly in the current literature.

The focus of this study is formulated into one main question: do the index traders have a negative impact on the roll return? In the current literature, it is found that activity of index traders may involve an increase of activity of speculators. Therefore, the secondary goal is to investigate whether speculators have an impact on the roll return. These questions are not investigated enough in the current literature and have practical implications for passive investors. The current literature predicts that the share of trader group positions to open interest may have an impact. The impact of the share of index trader positions and speculator positions is investigated in the thesis.

This study analyses the impact of index traders during the Goldman roll, because in this period, a majority of passive investors roll their positions and are exposed to this impact. While fundamental factors that drive commodity prices are not observed, many articles are impaired by not controlling for fundamental factors. In my thesis, I investigate and review fundamental factors that influence commodity futures prices and the term structure of commodity futures. The level of inventories, hedging pressure, open interest, the interest rate[5] adjusted for inflation and the yield spread are chosen on the basis of the current literature. After controlling for these fundamentals, I examine an impact of index trader positions on the spread. It is found that the main variables are non-stationary in levels. To examine their impact, the autoregressive distributed lag model is applied in this research. A wide range of commodities is analyzed in the paper. Seven agricultural commodities, four energy commodities, two livestock commodities and one metal commodity are included in the analysis. The publicly available reports of the CFTC are used to obtain data about positions of traders. Because Index Investment data (IID) of the CFTC is not available monthly before 2010, the OLS regressions on quarterly observations and implied Masters[6] positions are applied to forecast monthly data. To investigate the impact of speculators, the data of spreading positions of the CFTC is considered. To implement the front-running strategy speculators could open simultaneously short and long positions with different maturities of futures contracts. Such open positions are referred to as a spreading position in the CFTC reports.

The empirical tests confirm the negative impact of the share of index traders to open interest on the spread during the Goldman roll. The losses due to index trader activity per month may make up to 0.46% on average among commodities analyzed per month comparing the average changes of the nearest-to-maturity futures price of lean hogs equal to 0.56% over 2006 – 2013. This result measures the negative impact of the popularity of the GSCI and the DJ-UBSCI on returns of passive investors. For passive investors, this result implies that other indices would have less negative roll returns, until they do not roll a great number of positions in a short period of time. Recently introduced indices[7] that roll futures contracts in more backwardated, long-maturity futures suffer from a lack of liquidity of these futures. For less liquid futures, it is easier to achieve a high share of index traders, which distorts the roll return.

[5] 3-month Treasury Bill is chosen as a proxy for the interest rate.
[6] Details about forecasting and implied Masters positions are described in the Chapter “Data and descriptive statistics”.
[7] S&P GSCI Dynamic Roll http://us.spindices.com/indices/commodities/sp-gsci-dynamic-roll

It is also found that speculators push the spread down before the Goldman roll, however their impact is absent during the Goldman roll. Speculators push the spread down before the Goldman roll and push the spread up after the Goldman roll so that the total impact is not significant. The fact that speculators depress the spread before the Goldman roll also supports the negative impact of index traders during the Goldman roll. It is also discussed in the paper and is observed in the data of spreading positions that speculators do not open positions during or after the Goldman roll. It is difficult to use the arbitrage, because futures contracts are not liquid after the Goldman roll. This is because of the imminent expiration date. This result also infers that speculators do not push the spread to the fundamental value. For passive investors, it implies that investing in indices with non-disclosed rolling period[8] would not mitigate the problem of negative roll returns, which are affected by the share of index traders. It also infers that any limitation of activity of speculators would not have any impact on the roll return, if it does not influence the share of index trader positions.

My paper has the following structure: in the first chapter, I discuss the nature of index traders and the way of investing in commodities. The next chapter concerns the popular commodity indices. I focus on the fundamental factors that influence the spread in the third chapter. The fourth chapter concerns the related literature, both theoretical and empirical evidence of an impact of financialiazation on commodity prices. The fifth chapter discusses the impact of speculators on the spread. The sixth chapter deals with the data issues related to this work, methodology of calculating the index trader positions and descriptive statistics. The methodology is described in the seventh chapter. The implementation of methodology into practice and empirical results are discussed in the eighth chapter. The conclusion is presented in the ninth chapter.

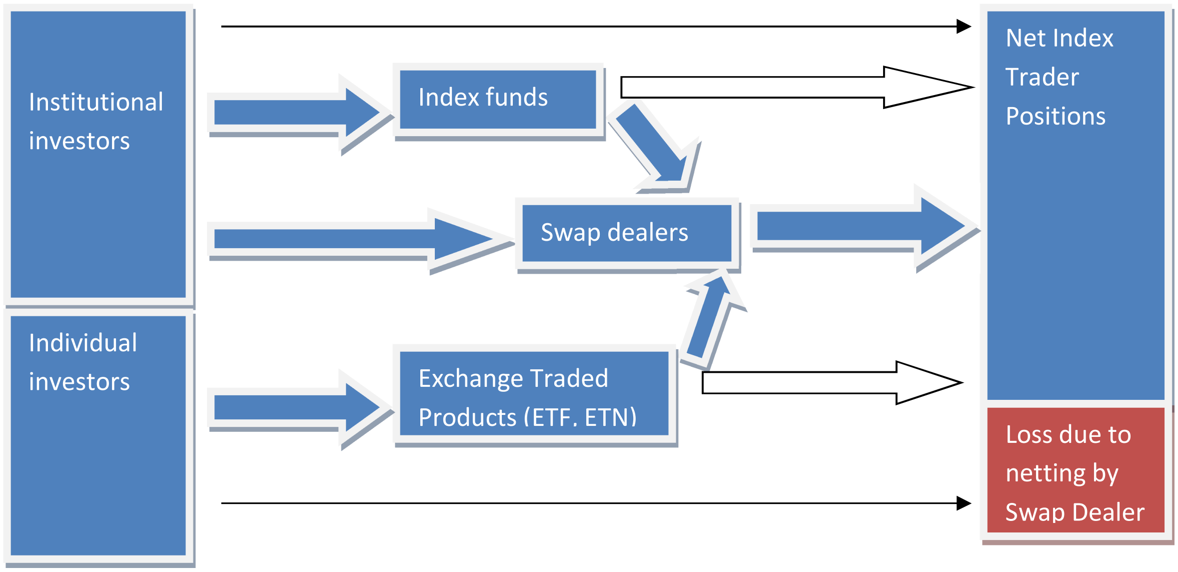
Index Traders

The deep and liquid futures market became popular among investors seeking benefits of diversification, because futures contracts synthetically generate returns of the spot market (Petzel, 2009). Historically, commodity futures returns have a low correlation with stocks and bonds (Anson, 1998). Furthermore, commodity futures returns have a significant correlation with unexpected inflation (Gorton & Rouwenhorst, Facts and Fantasies about Commodity Futures, 2006). However, trading costs were high before 2000s, that the expected returns from commodity investments were not able to outweigh the diversification benefits. The growth in trading volume of exchange-traded commodity futures made this market more liquid.

[8] For example, Deutsche bank liquid commodity index. http://en.wikipedia.org/wiki/Deutsche_Bank_Liquid_Commodity_Index

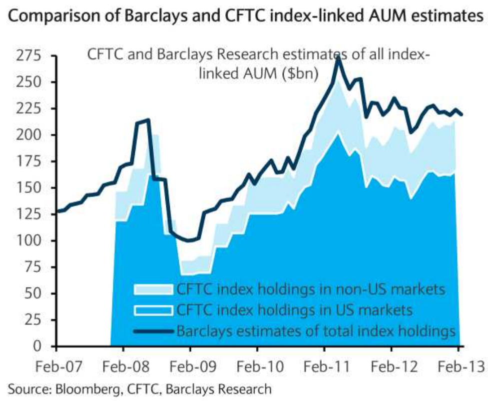
Figure 2. Flows of investment into the commodity market (Irwin & Sanders, Index Funds, Financialization, and Commodity Futures Markets, 2011)**.**

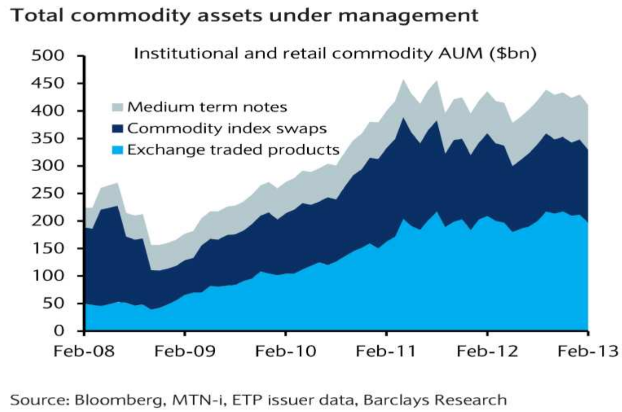
The thin arrows represent less used avenues for commodity index investment. The transparent arrows have lower impact in a comparison with non-transparent ones. Loss due to netting positions by swap dealers can be as little as 10% for agricultural commodities and quite large for energy and metals (CFTC, 2009).



To obtain higher diversification institutional investors invest in commodity index funds and commodity return swaps. With commodity index funds, institutional investors pool their commodity investment with a single fund manager and the manager agrees to manage the portfolio in a manner that mimics a commodity index benchmark. With OTC commodity swaps, institutional investors enter an agreement to receive the floating rate of return on a specified commodity index and paying the fixed rate over a specified period. The trading rules for index replication are transparent and well-defined, with pre-determined procedure of rolling of expiring futures contracts, rules of changing weights of commodities in the index. To capture the demand for commodities from individual investors the exchange traded products (ETFs, ETNs) were developed and has grown up dramatically since 2003 (Domanski & Heath, 2007). Shares of an exchange traded fund (ETFs) are trading on stock exchange and their price reflects the value of the index upon which it is based. Exchange traded note is a debt security which price is also linked to an underlying index. Figure 2 describes the direction of flows from investors to the commodity market (Irwin & Sanders, Index Funds, Financialization, and Commodity Futures Markets, 2011). Index funds, ETFs and ETNs can replicate the returns of an index by using futures contracts or by entering into swap contracts which provide with such returns, whichever is the most efficient. The largest flows go through OTC swap dealers (Irwin & Sanders, Index Funds, Financialization, and Commodity Futures Markets, 2011). The swap dealer may also have a customer who is a traditional short hedger. Hedger's position offsets the long position desired by an ETF. In this case, the swap dealer "nets" the two positions internally and may not have to go to the futures market to place a long hedge if the positions offset one another. If a swap dealer has many traditional long hedgers such as utilities, ETF positions would be smaller than index trader positions obtained after internal netting by swap dealer. According to the CFTC report (CFTC, 2009), swap dealer netting has been shown to be relatively small in the agricultural futures markets, it can be quite large for the energy and metals markets.

Figure 3. U.S. and non-U.S. commodity investments and total assets under management (Norrish & Croft, 2013).





In addition to the netting problem, there are issues related to counting U.S. versus non-U.S. investments. While the CFTC publishes regularly (since the recent time) data of commodity index trader positions, there is no relevant data about index trader positions for commodities trading on London Metal Exchange. Barclays Research regularly published data about their estimations of aggregate level of assets under management invested in commodities (Figure 3). However, it is published only annually. One can observe that the share of non-US index-linked assets under management to the total assets is not large on the aggregate level. It is worth noting that it is high for metal commodities that are not covered in my thesis except high-grade copper.

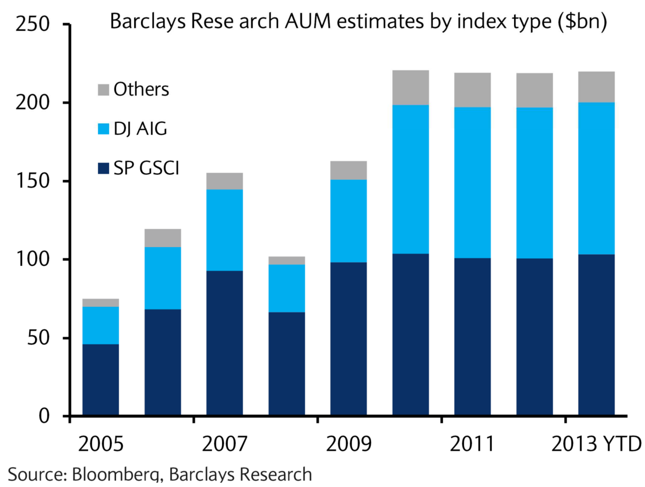
Commodity Indices

The most popular indices used as benchmarks for a majority of ETFs and ETNs are Goldman Sachs Commodity Index[9] (GSCI) and Dow-Jones UBS Commodity Index (DJ-UBSCI), formerly called DJ-AIG (initially introduced by AIG). They account for than 90% of assets under management linked to index trading of commodities. It is worth saying that the DJ-UBSCI share has risen during the last 5 years (Figure 4).

[9] Recently GSCI was acquired by Standard & Poor's, that's why sometimes it is referred as the S&P GSCI.

The GSCI was introduced in November of 1991. Its weights are based on the world production of the underlying commodities. The DJ-UBSCI was launched in July of 1998. The liquidity and world production determines the weights of commodities of this index. To avoid overexposure to any class of commodities there is a 33% maximum restriction for a weight. Likewise, there is also a 2% limit for a weight in the index. Both the GSCI and the DJ-UBSCI indices are well-diversified. The GSCI contains 24 commodities, versus the DJ-UBSCI including 19 commodities. The energy sector which has the largest weight in both indices accounts for 69% of the GSCI and 32% of the DJ-UBS (See weights of other classes in Appendix). Such a high share of the energy share in the GSCI is explained by the rules of construction. DJ-UBS have a maximum restriction, which leads to a lower weight of energy and a higher for other commodities. Because of a heavy weight to the one sector, over the period 2000 – 2009, the annualized standard deviation of the daily total returns of the GSCI was 25.9%, compared with 17.8% for the DJ-UBSCI (Irwin & Sanders, Index Funds, Financialization, and Commodity Futures Markets, 2011).

Figure 4. Share of the GSCI and the DJ-UBSCI indices in assets under management linked to index trading of commodities (Norrish & Croft, 2013).



The procedure of rolling futures contracts is also disclosed for both indices. Commodity indices usually hold contracts with short maturities, because the distant maturity contracts are not liquid enough (Mou, 2011). Different commodities have different rolling frequencies. The rolling scheme of the GSCI and DJ-UBSCI is displayed in Appendix (Appendix, Table 2). Agricultural commodities typically have 4-5 contracts per annum. The livestock commodities are rolled forward a bit more frequently, 6 to 8 times per annum. The rolling scheme of the GSCI is almost the same except for energy commodities, which are rolled every month versus every two month rolling of the DJ-UBSCI. According to the rules, the GSCI has a rolling period between the 5th and 9th business days of the expiration month. The DJ-UBSCI rolls from the 6th to 10th

business days, so the rolling periods of both indices greatly overlap (Mou, 2011). According to Mou, "many other indices and ETFs also roll in this period, like the former Lehman Brothers Commodity Index and the largest crude oil ETF: United States Oil Fund (USO). On each day of the rolling period, both indices roll forward 20% of the positions for commodities that need to be rolled". Brunetti and Reiffen use non-public data of positions of larger traders (Brunetti & Reiffen, 2011). They confirm for agricultural commodities that nearest-to-maturity contracts rolls to next nearest-to-maturity contracts in the period over 30 through 40 days to expiration.

Fundamental factors driving the spread

Before discussing the impact of index traders on the spread, the current literature about other factors influence the spread is surveyed in this paper. The following review covers three commodity specific factors. All three factors have both theoretical and empirical evidence about the impact on prices and the spread in particular. As for macroeconomic factors also reviewed below, most of them do not have a strong theoretical relationship. However, it has been found that some were significant in the past. Finally, some idiosyncratic effects are indicated as factors that are important for several commodities.

1) Inventories

The phenomenon of backwardation, when the spot price is lower than the futures price, has attracted the attention of researchers for a long time. Kaldor and Working brought forward the "Theory of Storage" (Kaldor, 1939; Working, 1949). Kaldor theorizes that the holder of the physical commodity receives a benefit that the holder of the futures contract does not have. The term of convenience yield that the holder of commodity earns has been introduced by Kaldor. Holder of physical commodity can take extra profits from temporary shortages in the market and is able to keep a production process running more efficiently. Later Deaton and Laroque have extended the model by the theoretical explanation of a link between inventory level and future spot price volatility (Deaton & Laroque, 1992). A holder of physical commodities uses inventories as against supply and demand shocks. So the inventory level reflects the expectations about future balance of supply and demand. The empirical test of this theory has been carried out by Gorton (Gorton, Hayashi, & Rouwenhorst, The Fundamentals of Commodity Futures Returns, 2007). The relationship between the spread which is used as a proxy for the basis[10] and the inventory level is directly examined for a large cross-section of commodities. The authors show

[10] In the academic literature, the spread between futures and spot prices called as the basis. However, the data about the spot price is limited. As a result, the difference between two nearest-to-maturity futures is used as a proxy for the basis.

that the inventory level has a negative relationship with the futures basis. That means that lower inventory level leads to a higher backwardation of commodity prices. In addition, it has discovered that discretionary (adjusted to trend and seasonality) level of inventories has a strong predictability of commodity futures returns (Dincerler, Khokher, & Simin, 2006). *To sum up, I conclude that inventories have a negative impact on the spread.*

2) Hedging Pressure

At the same time with the "Theory of Storage", the "Theory of Normal Backwardation" had been developing by Keynes (Keynes, 1930). This theory argues that the difference between spot and futures price depends on the balance between hedgers that take short positions and speculators that trade against hedgers. When a number of hedger's positions exceed a number of speculators, the futures price goes down and the term structure becomes more backwardated. This theory is empirically supported by several articles (de Roon, Nijman, & Veld, 2000; Basu & Miffre, 2013). However, it has been found that the impact is less significant for the closest to maturity commodity futures (de Roon, Nijman, & Veld, 2000). Besides Gorton and Rouwenhorst notice that positions of hedgers are contemporaneously correlated with inventories and futures prices and there is no evidence that these positions are correlated with ex-ante risk premiums of commodity futures (Gorton & Rouwenhorst, Facts and Fantasies about Commodity Futures, 2006). The recent article of Acharya (Acharya, Lochstoer, & Ramadorai, 2013) points out that the net hedging pressure from the CFTC data[11] has significant shortcomings. "The traders that have a cash position in the underlying can obtain a hedger classification. This includes consumers of the commodity, and more prominently, banks that have offsetting positions in the commodity (perhaps on account of holding a position in the swap market). The line between a hedge trade and a speculative trade, as defined by this measure, is therefore blurred". Acharya uses 2 different measures of default risk of producers that can incline the producers to hedge. It has been concluded that the aggregate hedging activity depresses the futures prices, but in the absence of an inventory stock-out hedgers also decrease inventory holdings which in turn decreases the spot price. The argument of Acharya about decreasing the spot price weakens the impact of hedging pressure in a long run. *To summarize, a positive contemporaneous impact on the spread is anticipated because of the hedging pressure.*

[11] U.S. Commodity Futures Trading Commission (CFTC) publishes weekly data about positions of commercial and non-commercial traders on the U.S. commodity exchanges.

3) Growth of total open Interest

Hong and Yogo make an empirical and theoretical evidence that open interest contains information about future economic activity and inflation expectations that is not fully revealed by futures prices or net supply-demand imbalances among hedgers and speculators (Hong & Yogo, What does Futures Market Interest Tell Us about the Macroeconomy and Asset Prices?, 2012). The authors notice that total open interest tends to be pro-cyclical. Anticipation of higher demand forces producers to take a short position and utilities to take a long position on futures market. Under assumption of underreaction of non-informed speculators and limited risk absorption capacity, open interest commoves with futures price. The empirical test reveals that the 12-month geometrically averaged growth of open interest is positively correlated with the monthly excess returns of fully collateralized commodity futures. Additionally, Sockin and Xiong suggest a consistent idea that in the presence of a high complementarity an increase of open interest is associated with the expectations of consumers of commodities (utilities) about the higher global demand. *As a result, the price of futures increases, which implies higher contango of the term structure.*

4) Macroeconomic factors

Several articles have examined the impact of the macroeconomic factors on commodity prices and the term structure. The real interest rate is the component of the storage cost according to the “Theory of Storage”. As a result, a higher interest rate leads to higher relative futures price or higher contango. In addition, Frankel has carried out tests and has found the empirical evidence of this relationship (Frankel, 2006). The author also argue that “high real interest rates reduce the demand for storable commodities, or increase the supply, through a variety of channels: 1) by increasing the incentive for extraction today rather than tomorrow; 2) by decreasing firms' desire to carry inventories; 3) by encouraging speculators to shift out of commodity contracts (especially spot contracts), and into treasury bills”. On the longer term (second order) there is opposite effect on the spread, because higher real interest rates lower the inventories level. This makes the curve more backwardated.

Fama and French notice that the term structure of industrial metals became backwardated during the business-cycle peaks of 1973-1974 and 1979-1980 (Fama & French, 1988). Early and recent articles have observed the positive correlation of the yield spread (a difference between Moody’s AAA corporate yield and the short rate) and the backwardation of commodities (Hong & Yogo, Digging into Commodities, 2009). Dempster, Medova and Tang consider a difference between 10-year T-bond and 3-month T-Bill as a proxy for recession. They have found also a positive correlation with the convenience yield (Dempster, Medova, & Tang, 2012). The usual interpretation is that the yield spread can be a proxy for recessions (Estrella & Mishkin, 1998). Hamilton notices that 10 out of 11 recessions were preceded by the sharp increase of spot prices (Hamilton, Nonlinearities and the Macroeconomic effects of Oil Prices, 2011).

Bailey and Chan have investigated the impact of the corporate spread on the futures basis (Bailey & Chan, 1993). The corporate spread is defined as a difference between the annualized yield of a low-grade corporate bond portfolio (with a Moody Rating of BAA) and the annualized yield of a high-grade corporate bond portfolio (with a Moody Rating of AAA). According to Bailey and Chan, while interest rate is a component of the cost of storage, the corporate spread can capture the extra risk premium that "reflects the systematic risk of the underlying commodity". The corporate yield is a countercyclical: it is low in times of high economic activity. The authors suggest that the predictive power of the corporate yield is due to the sensitivity of the basis to the corporate default risk, for which the corporate yield is used as a proxy. This idea is corresponded with Acharya's paper (Acharya, Lochstoer, & Ramadorai, 2013) that shows that the corporate default risk inclines the producers to hedge more. Therefore, extra hedging pressure due to corporate default risk should be positively correlated with backwardation of the term structure. However, Bailey and Chan have observed the significant impact of this variable with different signs for different commodities.

Bailey and Chan also detect a positive correlation of the dividend yield with the backwardation of the term structure. But recent articles note that the effect of the dividend yield is insignificant (Hong & Yogo, Digging into Commodities, 2009). Moreover, the effect of the dividend yield does not have an appropriate interpretation for the commodity market. The index of the implied volatility of options (VIX) is also examined as a factor that drives the term structure. However, Hong and Yogo have found that it has a weak significance. The currency exchange rates (especially the dollar rate) have been also investigated for explanation of the variability of futures prices (Akram, 2009). According to Akram, "a decline in the value of the dollar may be due to a reduction in the US interest rates. Uncovered interest rate parity implies that the expected depreciation of the dollar would be inversely related to the interest rate spread between the US and abroad. It follows that a reduction in interest rates may both directly and indirectly via the depreciation of the dollar lead to higher commodity prices". Dempster, Medova and Tang examine the dollar index growth and show that it has a weak impact on the convenience yield[12].

[12] Dempster, Medova & Tang (2012) estimated the impact of the dollar index on the 3 latent factors that are associated with the short-term, medium-term and long-term convenience yields by using the Structural Autoregression model.

Due to the fact that there is a poor evidence of an impact of the dividend yield, the VIX and the dollar index growth (Hong & Yogo, Digging into Commodities, 2009), these factors are lacking in an appropriate interpretation and are not included in my analysis. Initially I have included the VIX, but the estimated coefficients of the VIX in the regressions turns out to have an inconsistent sign among different commodities and frequently not significant even at 10% level. The corporate yield is highly correlated with the hedging pressure and associated with default risk, which is also captured by the hedging (Acharya, Lochstoer, & Ramadorai, 2013). The factors included in my analysis are present in Table 1.

Table 1. The anticipated impact of the factors on the spread.

Factors	Inventory level	Hedging Pressure	Open interest growth	Real interest rate	Yield spread
Anticipated sign of marginal effect	negative	positive	negative	negative	positive

5) Idiosyncratic effects

In addition to the variables discussed above, an impact of idiosyncratic factors has been investigated in the past. Unpredicted climate and weather effects have a significant impact on prices of many commodities (Mu, 2007). Roll (Roll, 1984) shows that there is a statistically significant relation between orange juice returns and subsequent errors in temperature forecasts. Bruner (Brunner, 2002) have found the considerable impact of the hurricane on real commodity prices. The changes of market regulation influence volatility of futures prices (Dong, Du, & Gould, 2011). Political events in OPEC countries influence prices of crude oil (Kilian, 2009). Although these factors account for a significant variation of prices of many commodities, these factors are hardly measured and observed.

According to this review of the literature, the anticipated impact of the factors is illustrated by Table 1. The negative sign of marginal effect means that positive changes of the factor relate to the negative changes of the spread. For example, an increase of the inventory level leads to lower spread or higher contango. The impact of all these factors has been verified in the recent literature. However, there is no paper that examines them all together. All these factors are examined as control variables to check whether index traders affect the spread of commodity prices.

Index trader impact on commodity prices

Theoretical evidence

To avoid delivery of commodities the index trader positions are rolled before expiration of the future contract to the next nearest to maturity futures. The proponents of the negative impact of financialization on the spread argue that the relative size of index trader positions is huge to have an impact on the spread. The index trader could depress the nearest-to-maturity futures and inflate the next nearest to maturity futures. One of the articles discusses the impact of index traders is the paper of Brunetti and Reiffen (Brunetti & Reiffen, 2011). The authors develop the theoretical model and claim that the model relates to the effect of index trader positions on the term structure of futures contracts. The model predicts that, the spread between the first deferred and the nearby contract depends on both the relative sizes of commodity index trader (CIT) positions in the two maturities, and on the aggregate size of CIT positions, albeit in a way that varies across the contract. This model is consistent with the equity market. The authors assert that the redefinition of the index leads to an increase in the demand for the futures contract, as some mutual funds are contractually obligated to have a portfolio that is representative of a specific index. Thus, the increased price results from a liquidity effect: i.e. the interaction of this higher demand with a less-than-perfectly elastic supply. The implication of their model is that the cost of hedging falls as index trader positions increase.

Several recent articles have discussed the theoretical evidence of an impact of index traders on correlation among index included commodities and volatility of spot and futures price. In the article of Basak and Pavlova (Basak & Pavlova, A Model of Financialization of Commodities, 2013) the dynamical equilibrium model with both institutional investors and standard market participants has been developed. The main idea is that the institutional investors care about their performance relative to a commodity index. The authors postulate that “the marginal utility of institutional investors increases with the index. Both classes of investors in the model invest in the commodity futures markets and the stock market”. One of the implications of the model is that if commodity prices go up, the index included commodity prices go up higher. The model allows disentangling the effects of financialization from the effects of demand and supply (“fundamentals”). They perform a simple calibration and figure out that the financialization explains for 11% to 17% of commodity futures prices and the rest is attributable to fundamentals. Additionally, they found that “if a commodity is included in the index, supply and demand shocks specific to that commodity spill over to all other commodity futures markets. In contrast, supply and demand shocks to a nonindex commodity affect just that commodity market alone." The authors also observed that the financialization amplifies the effect of demand and supply shocks. For example, they estimated that "a 33% increase in demand for a commodity raises the fraction of its futures price attributable to financialization from 16.8% to 24.9%". The similar model and results were obtained by Baker (Baker, 2012).

The impact of passive investors on spot prices has been also investigated in literature. Hamilton (2009) shows that if the demand is perfectly inelastic and speculators force up futures prices, this would force spot prices to maintain equilibrium in the storage market. Petzel notes that futures positions are synthetic positions in the spot market (Petzel, 2009). So an increase of futures prices leads to higher spot prices. Sockin and Xiong (Sockin & Xiong, 2012) have found a correlation between open interest and futures price; they have interpreted this as a signal for higher growth. The authors also found that higher open interest leads to higher spot price as a result of "feedback effect". Baker (Baker, 2012) deems that the effect of index traders on spot price is weak.

Empirical evidence

While a majority of articles investigates the impact of index traders on commodity prices, futures return and correlation of returns. There are a few articles that discuss the influence of index traders on the spread or the basis of commodities. Moreover, even a fewer number of articles examines the impact of index traders during the rolling period. One of such articles, the article of Brunetti and Reiffen (Brunetti & Reiffen, 2011) analyses the daily variation of the spread. As stated above, Brunetti and Reiffen examine the impact of index traders on spread and made some theoretical findings. In addition, they implement empirical tests which confirm their ideas: "the spread falls as CITs move their positions from the nearby to the first deferred". They analyse only three agricultural commodities, namely soybeans, wheat and corn, over the period 2003 – 2008. Brunetti and Reiffen use non-publicly available data from the US Commodity Futures Trading Commission Large Trader Reporting System (LTRS). This database contains information about all end-of-day positions of each large trader. According to the authors, such large traders represent 70-75% of all open positions. Such database allows disaggregation between individual traders and between different maturities. It allows them to test the relationship between day-to-day changes in CIT positions and the associated price changes, in addition to test the effect of CIT positions on the spread with the harvest cycle. They detect that the effect of the CIT positions is larger just prior to the harvest.

The current literature researches an "anomaly" of futures return during the roll period. The presence of this phenomenon confirms the negative impact of index traders during the rolling period. Stoll and Whaley (Stoll & Whaley, 2010) assert that the returns are more positive but not significantly during the rolling period in a comparison with the rest period. But the authors observe the positive significant results for crude oil in a rolling period. Mou (Mou, 2011) also have investigated the impact of rolling of index traders on the term structure of different commodities. The author examines the strategy which trades in the opposite direction to the index traders. With very simple front-running strategies, the arbitrageurs can "generate excess returns with positive skewness and annual Sharpe ratios as high as 4.39 over the period from January 2000 to March 2010". The profitability of the strategies is positively correlated with the net result of two opposite forces: the size of index investment and the amount of arbitrage capital employed. The arbitrage profit is lower when there is a reduction in index investment or an increase in arbitrage capital.

The rest part of the literature does not differentiate the periods when the Goldman roll happens and when it does not. Instead, it considers the impact of index traders over the whole period of observations. I divide the spectrum of the literature about the financialization of commodities into three parts by using different methodology and investigating different impact of index traders.

In the first part researchers apply the Granger causality test to examine whether the lagged changes of index traders positions help to forecast futures returns and vice versa. Gilbert (Gilbert, 2009) has found the confirmation of significant causality of index trader positions on returns for 3[13] out of 7 commodity markets investigated. However, several articles have observed that there is no significant impact of index trader positions on futures returns (Irwin & Sanders, Index Funds, Financialization, and Commodity Futures Markets, 2011; Stoll & Whaley, 2010). It is worth saying that application of granger causality test has several well-known flaws. In details the limitation of the granger causality test is discussed in Grosche's paper (Grosche, 2012). By construction the test does not allow to examine the contemporaneous influence. In addition to that, other flaws originate from omission of relevant variables, quality of data, forward-looking behavior[14] and time-varying effects. The fact that the number of monthly observations is very small, while weekly futures return are very volatile cause the application of the granger causality test even more difficult.

The second spectrum covers the literature about an increase of correlation between index included commodities and between index included commodities and equity market. The same articles also consider the effect of index traders on volatility of futures prices. Silvennoinen and

[13] Namely crude oil, aluminum and copper.
[14] Forward-looking behavior means that the current index trader positions contain the information about the expectation of return in a future. Such behavior leads to failure to reject the hypothesis that one variable cause the another one.

Thorp (Silvennoinen & Thorp, 2012) have observed the higher correlation between commodity and stock market returns over the recent years. Tang and Xiong (Tang & Xiong, 2010) show that the correlation between non-energy and crude oil futures returns has increased after 2004. Additionally, while the correlation between non-index commodity returns grew up to 0.2 after 2004, the correlation between in-index commodity returns climbs up to 0.5. The authors point out that such a high correlation has not been observed over 1975 to 2004. In both articles (Silvennoinen & Thorp, 2012; Tang & Xiong, 2010) the evidence of volatility spillover from equity market to commodity market has been observed. Stoll and Whaley also analyze the comovement between in-index and off-index commodity prices. They make an opposite conclusion for 2006 – 2009 observations. Several articles focus on the specific commodity market, for example on the cotton market (Janzen, Smith, & Carter, 2012). According to the authors, the cotton was a single commodity that did not have a lack of supply and all fundamentals were stable in 2006-2008. However, the price was also volatile and the comovement with other commodities has increased. The Structural Vector Autoregression (SVAR)[15] is introduced by the authors to explain such behavior. The authors find a weak impact of financialiazation.

The third layer of literature examines the direct impact of index traders on futures return by using linear regressions. Singleton (Singleton, 2011) shows that after controlling for open interest growth, emerging market growth, U.S. stock returns and lagged futures return, the index trader positions' growth has a positive significant effect on futures return of oil.

In the end of this review of the literature I would like to discuss some arguments of Irwin and Sanders (Irwin & Sanders, Index Funds, Financialization, and Commodity Futures Markets, 2011) against the significant relationship of the index trader positions and futures prices. The first argument is that speculative needs did not exceed hedging needs since 2006. The authors make a conclusion that is the reason why fundamentals drive the prices. However, this argument does not reject that financialization may produce an impact on futures prices weaker than fundamentals do. The authors also claim that price movements in future markets with the substantial index fund performance were not uniformly upward in the spring of 2008. This argument also supports the idea that fundamentals are the main driver of prices, but it does not reject a limited impact of index traders. The third claim they raise is that high prices are also observed in commodity markets not connected with index fund investment. Tang and Xiong (Tang & Xiong, 2010) also observe this case, but they notice that the correlation among in-index

[15] It is worth saying that the SVAR models have similar flaws with the granger causality tests. In addition, the problem of identification can be a crucial issue, which defines the results. http://www.ifw-members.ifw-kiel.de/publications/an-introduction-into-the-svar-methodology-identification-interpretation-and-limitations-of-svar-models/kap1072.pdf

commodities is much higher. Irwin and Sanders also contend that the “noise traders” can increase prices only if they are not predicted, but index positions are predicted. However, there is still some uncertainty associated with inflows and outflows from index funds. In addition, the activity of speculators that front-run the Goldman roll is shown to be significant (Frenk & Turbeville, 2011). Their activity also adds uncertainty during the rolling process.

Impact of speculators

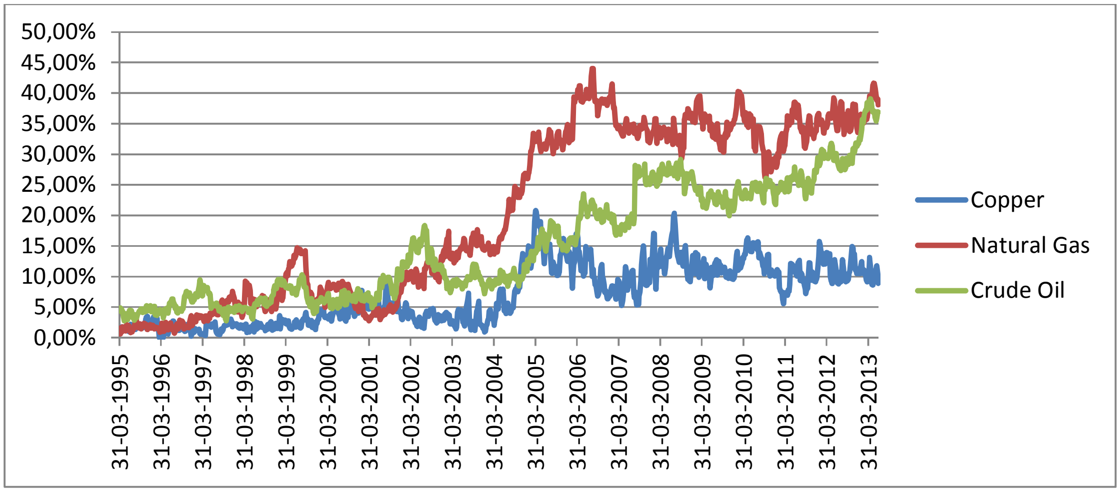
As stated above, the rolling procedure is publicly available. More than 95% of index trader positions are associated with the GSCI and DJ-UBSCI. All of these positions are rolled in the predetermined dates. Therefore, the impact of index traders should be mitigated by speculators (or rational news traders) according to Friedman (Friedman, 1953). Speculators may take the opposite positions to index traders: buy nearest-to-maturity futures and sell next nearest-to-maturity futures during the rolling process of index traders. However, after the rolling of index traders, only few weeks are left before expiration. Because of a lack of liquidity in the period close to expiration, it would be very expensive to close a considerable number of positions. As a result, the implementation of such strategy is difficult for a traditional speculator. At the same time, such strategy can be accomplished by speculators that can afford the delivery of commodities in warehouses. There is some evidence that the largest institutional investors[16][17] and hedge funds[18] hold physical commodities in the exchange warehouses. These facts indicate indirectly about possible implementation of such a strategy.

Yet, speculators may exacerbate the impact of index traders. Long et al (Long, Shleifer, Summers, & Waldmann, 1990) propose a theoretical model showing that “rational news traders, by anticipating the price impact of trend followers (or positive feedback traders), actually end up destabilizing markets. In their model, rational speculators, in anticipation of the forthcoming buy/sell orders of trend followers, increase their long/short positions today in the hope of earning higher returns tomorrow”. As a result, instead of stabilizing prices, they end up setting price trends and deterring short-term prices away from fundamentals. The profits of this strategy have been investigated by Mou (Mou, 2011). The author examines the strategy that takes a short position on the nearest-to-maturity futures and takes a long position on the next nearest-to-maturity futures several days before the rolling of the GSCI (“Goldman roll”) starts. These positions are closed 10 days after. It is found that such strategy yielded excess returns with positive skewness and Sharpe ratios as high as 4.39 from 2000 to 2010. However, the author

[16] http://ftalphaville.ft.com/2011/05/24/576501/the-uk-is-concerned-about-banks-that-warehouse-commodities/
[17] http://online.wsj.com/article/SB10001424052702304803104576426131256469252.html
[18] http://www.ft.com/intl/cms/s/0/41c44bc2-e107-11de-af7a-00144feab49a.html#axzz2YBpR3Opz

notices that the profits of this strategy have dropped dramatically in a recent period. According to the author, the arbitrageurs became more aware about this trading opportunity after 2004 and more capital has been utilized to exploit this market anomaly. As a result, profits almost disappear.

Figure 5. The ratio of spreading positions and open interest of Copper, Natural Gas, Crude Oil, Datastream, CFTC.



Frenk and Turbeville (Frenk & Turbeville, 2011) point out that a number of spreading positions of non-commercial traders has increased after 2006. “Spreading” is a computed amount equal to offsetting long and short positions held by a trader. The computed amount of spreading is calculated as the amount of offsetting futures in different calendar months or offsetting futures and options in the same or different calendar months. One can observe a massive growth of spreading positions after 2004 for three commodity markets in Figure 4. In Appendix, the graph of spreading positions of two agricultural commodities included in the index and two non-index commodities is presented. The graph of spreading positions of in-index commodities has an ascending trend after 2004 and their share has achieved more than 20% of open interest. At the same time, the non-index commodities do not have a trend after 2004 and the share of spreading positions is close to zero. These graphs confirm the idea that these spreading positions are used for the implementation of front-running strategy for in-index commodities.

According to Frenk and Turbeville, this capital inflow from speculators does not bring about lessening of contango during the rolling period. Instead, speculators push the spread down before the rolling period by applying the front-running strategy. Index traders are insensitive to prices of futures contracts and roll their positions irrespectively on activity of other traders. As a result, during the rolling period, the spread drops further. The authors notice that there are no observation techniques that allow market participants to discern between trading-generated signals and supply and demand generated signals. The futures market provides the information that is derived from an artificial financial source. This information becomes misinterpreted by many market participants as a sign that higher prices are due to commodities fundamentals, rather than due to index trader activities. As a result, the impact of index traders is higher during the periods of supply and demand shocks.

Hypothesis development

A majority of articles studies impacts of pressure from a trader group on commodity prices focuses on net positions held by a trader group (de Roon, Nijman, & Veld, 2000, Buyuksahin & Robe, 2011). Brunetti & Reiffen (2010) develop theoretical and empirical evidence of the negative impact of index trader positions on the spread. The authors examine the impact of the share of index traders to the open interest on the spread. The first hypothesis I test is:

The share of net index trader positions to open interest has a negative impact on the spread of commodity futures during the rolling period.

The confirmation of this hypothesis would imply that the rolling of huge amount of index trader positions decreases *t*he spread. As a result, it decreases the returns of passive investors. The confirmation also implies that, in order to minimize the detrimental impact of the rolling process index, traders should avoid rolling in futures and periods when the share of their positions is high. The rejection of the hypothesis would imply that the share of positions doesn't have any impact on the roll return, which is explained by fundamental factors.

Mou (Mou, 2011) points out that the number of spreading positions of speculators has surged after 2006. Several articles claim that the speculators implement the front-running strategy, which then distorts the price before the rolling period (Mou, 2011; Frenk & Turbeville, 2011). The fact of implementing this strategy implies that it is profitable for speculators. It also confirms that index traders have a negative impact on the spread during the rolling. The second hypothesis I test is:

The share of spreading positions of speculators to open interest has a negative impact on the spread of commodity futures ***before the rolling period****.*

If the hypothesis is confirmed, it would imply that speculators open spreading positions that push the spread down. It also supports that index traders push the spread down during the rolling. The rejection of the hypothesis does not contradict the confirmation of the first hypothesis. However, it means that speculators do not influence the spread before the rolling period.

Frenk and Turbeville argue that speculators may have a negative impact on the spread and push it to lower values. However, there is no evidence whether the speculators drive the spread down and exacerbate the negative effect of index traders or they increase the spread and lessen the effect of index traders. Mou observes that the profits from the front-running strategy have dropped after 2006. However, this result does not imply directly any direction or absence of their impact. The profits might decrease, because speculators push the spread up during the rolling period. At the same time, a decrease of profits of speculators might imply that they have started to open spreading positions earlier than Mou assumes in the tests. In order to understand whether speculators push the spread to lower values than index traders do during the rolling period, the third hypothesis is checked in the thesis:

The share of spreading positions of speculators to open interest has no impact on the spread of commodity futures under controlling for the share of net index trader positions ***during the rolling period****.*

The confirmation of this hypothesis would imply that speculators do not push the spread up or down during the rolling period if the share of net index trader positions is fixed. At the same time, this result does not contradict with the confirmation of the second hypothesis. In this case, it would imply that speculators push the spread down before the rolling period and after the rolling period they push the spread up with the same magnitude. This means that the total impact on the roll return disappears during the rolling period. If the third hypothesis is rejected, the implication would depend on the sign of the impact of the share of spreading positions. In the case it is positive, it would mean that speculators overtrade and push the spread to fundamental values. If it is negative, it would infer that speculators drive the spread to lower values than index traders do. The impact of the behavior of speculators has an implication for agencies that regulate the commodity market and may restrict the activity of speculators. For example, they may limit the size of positions. In the end, I would like to point out that my paper avoids the discussion of the impact of the spread in the non-roll periods[19]. The analysis of these periods implies also the long-term impact on the spread and spot prices, which is under heavy debate in the current literature. At the same time, if this impact is present, a majority of articles presume that it should be weak. If it is weak, it should not distort my estimation of the effect during the rolling period, which is important for passive investors. Ultimately, I leave this topic for further research.

[19] Excluding a short period before the Goldman roll, which is examined for testing the impact of speculators.

Data and descriptive statistics

The spread of commodities

The data about futures price of the first two nearest-to-maturity contracts is available in Thompson Reuters Datastream. The analysis of prices of Sweet Light Crude Oil, Heating Oil, Gasoline[20] and Natural Gas that are trading on the New York Mercantile Exchange is carried out in this paper. Among commodities trading on the Intercontinental Exchange in New York coffee "C", cocoa, cotton, sugar n. 11 are captured in this research. Copper trading on the COMEX is also included. Such commodities trading in the Chicago Board of Trade as wheat, soybeans and corn are involved as well. Finally, prices of lean hogs and live cattle trading in the Chicago Mercantile Exchange are taken into consideration. The expiration dates of futures contracts for all commodities are present in Appendix[21]. Consequently, one can estimate the spread at time t:

$$Spread_t = \frac{\log(F(t,T)) - \log F(t,T+k)}{T-t} \qquad (1)$$

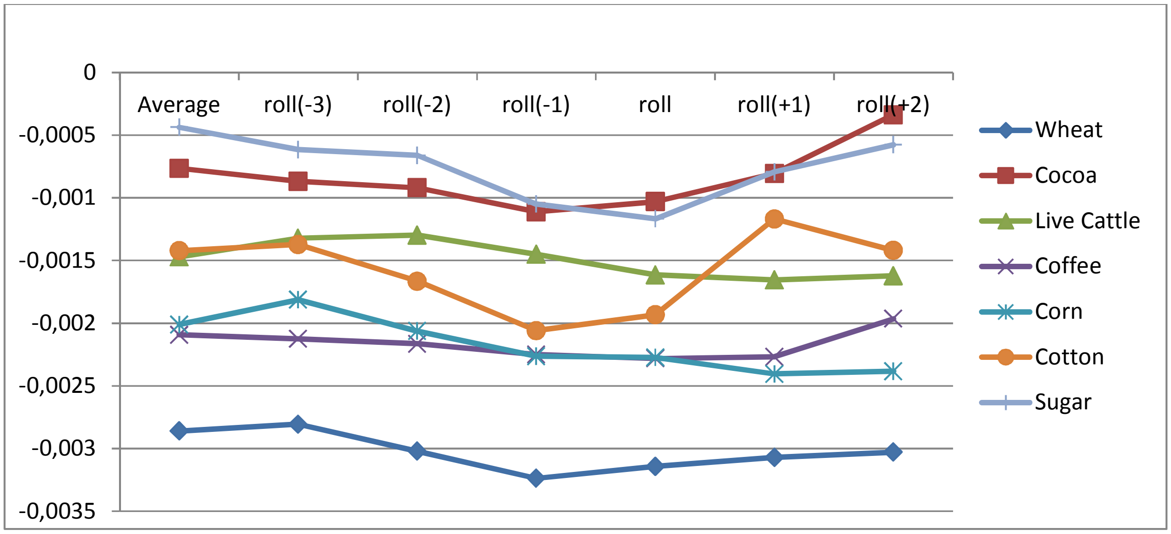
T is a date of an expiration of the first nearest-to-maturity futures contract, $T+k$ is a date of an expiration of the second nearest-to-maturity futures contract. Such way of calculating spread corresponded with a profit of a trader due to the rolling process. During the rolling process a trader takes a short (sell) position for the nearest futures contract and takes a long (buy) position to the next futures contract. If the spread is positive, an investor earns on the rolling process. In a case of the negative spread, she sustains losses.

The summary statistics about the spread of 14 commodities taken into consideration is present in Appendix. Table 3 of Appendix contains the statistics of the spread for a period when the data of index trader positions[22] is present for each commodity. Table 4 shows the statistics over a long period from 1983 to 2013 depending on availability of data. According to these tables, all agricultural commodities, crude oil and natural gas are in a contango on average, because their mean is negative. At the same time, heating oil, gasoline and copper on average in a backwardation. One can observe that a majority of these commodity markets have a lower spread in a short sample. However, Gasoline, Cocoa and Sugar showed more backwardated term structure in a recent time: the spread increased in a recent time compared with their average value. Noteworthy, the spread is the most volatile for livestock commodities and natural gas.

[20] New York Harbor RBOB Gasoline.
[21] The exact expiration dates and rules are possible to read on the sites of the corresponding commodity exchanges. The current expiration dates are available in the Capitol Commodity Services site. For example, for sugar: http://www.ccstrade.com/futures/SB/exp/
[22] The details about index trader positions are present in the next chapter.

Figure 6. The average spread in the Goldman roll period and over three weeks before to two weeks after the Goldman roll.

"Average" is the average spread over the whole period after 2004, Datastream, weekly data



.

To illustrate the impact of index traders and speculators applied front-running strategy, I calculated the average spreads in weeks when Goldman rolling happens and in weeks before and after the Goldman roll[23]. Figure 6 shows that the spread goes down at least one week before the Goldman roll. For all commodities except Live Cattle the spread starts to decrease in two weeks before the Goldman roll. In all cases the spread is lower than the average spread over the whole period after 2004. In Appendix I also included the same graph but for observations before 2004. The behavior of the spread moves differs from the observations after 2004. One can see that the spread is higher during the considered period than the average spread for all commodities except cocoa and wheat. The spread is the lowest for weeks after rolling what can be explained by seasonality factors.

Fundamental factors

The total open interest and information about commercial and non-commercial positions are obtained from DataStream as well. These data is collected by the CFTC and reported in the legacy commitment of traders (COT) reports. Hedging pressure is calculated in the following way:

$$Hedging\_Pressure_t = \frac{Commercials\_Short_t - Commercials\_Long_t}{Commercials\_Short_t + Commercials\_Long_t} \quad (2)$$

[23] I calculated the roll period spreads only for agricultural commodities. Energy commodities and copper (COMEX) are rolled every month. Datastream joins (roll) futures price series in the end of the month. As a result, continuous time series of futures prices does not contain information about weeks before the rolling (more than one week).

where *Commercials_Short* is a total number of short positions hold by commercials, *Commercials_Long* is the same of long positions. These data is also collected by the COT reports of the CFTC (the details about classification below and in Appendix). The data about commercial and non-commercial positions is absent for metals trading in the LME. Furthermore, futures prices of the nearest futures contracts are not available as well. For these reasons, I do not consider these commodities in my research.

To get the data of inventories I followed the sources of Gorton's paper (Gorton, Hayashi, & Rouwenhorst, The Fundamentals of Commodity Futures Returns, 2007). However, I do not have access to Commodity Research Bureau database. As a result, I used alternative sources of data about inventories such as the Bloomberg database (Appendix, Table 6). Nevertheless, I was not able to find information about inventories for commodities such as soybean oil, soybean meal and feeder cattle.

As for macroeconomic factors, the 3-month T-bill adjusted is considered a proxy for the nominal rate. The twelve-month growth of the CPI is utilized as a current rate of inflation. The difference between two variables is assumed as a proxy for the real interest rate. To capture effect of recessions on commodity prices, I used a difference between the 10-year Treasury bond rate and the 3-month federal funds rate, which has been shown to be a good proxy for recessions (Estrella & Mishkin, 1998). All these variables are available in Thompson Reuters Datastream.

Index Trader Positions

Each Tuesday the CFTC publishes the legacy commitment of traders (COT) reports. These reports contain the information about the number of contracts held by commercial and non-commercial traders that are above a reporting level. These reports with the data from 1993 for a majority of commodities can be accessed by DataStream by Thompson Reuters. The earlier research used this data to examine the role of speculation in the commodity market. However, index traders invest the most via swap dealers. As a result, the number of non-commercial positions of COT is not appropriate measure for a number of index trader positions. To keep a track of level of swap dealer positions the CFTC released a new set of data that broke down the two parts into subcategories. The subcategories of the Disaggregated COT reports are shown in Appendix, Table 1. Although the Disaggregated COT reports split the COT reports, it is still impossible to obtain the exact number of index trader positions. To capture the index trader positions the CFTC decided to report the Supplemental COT. Supplemental reports show aggregate futures and option positions of Noncommercial, Commercial, and Commodity Index Traders (CIT) in 12 selected agricultural commodities. The decision to classify a trader as an index trader was made according to the analysis of positions and series of interviews[24]. However, these data is exposed to preponderance of a trader to one strategy. If a trader has a majority of positions replicating indices, all her positions would be considered as index trader positions. Moreover, classification of an index trader is defined by her "ex-poste" trading pattern. Changes of strategy are not considered. The changes of trading pattern and a presence of considerable part non-index trading positions are typical for a swap dealer. However, for the agricultural commodities it is not a great problem, compared with the energy commodities (Irwin & Sanders, Index Funds, Financialization, and Commodity Futures Markets, 2011). These weekly reports are accessible in DataStream since January of 2006. To solve the problem with the preponderance of the trader's trading strategy the CFTC organized the "special call" to swap dealers and index funds. The so-called Index Investment Data (IID) is considered the best data for index trader positions[25] (Irwin & Sanders, Testing the Masters Hypothesis in Commodity Futures Markets, 2011). However, the CFTC publishes only monthly IID reports for 21 commodities since June of 2010 and quarterly reports since December 2007.

As stated above, Supplemental COT had the data only about 12 agricultural commodities. Masters (Masters & White, 2008) suggests a popular procedure how to estimate index investments for other commodities. The example of how to estimate the index trader positions for crude oil is present in Appendix (p. 34). Masters assumes that the GSCI and the DJ-UBSCI[26] are the only two commodity indices. He noticed that feeder cattle and Kansas wheat are unique commodities that are present in the GSCI. At the same time, soybean oil is unique for the DJ-UBSCI. To calculate the size of the two indices, one can multiply a number of contracts invested in commodities by a contract size and then divide by a weight in the index. Irwin and Sanders argue that such procedure can lead to the results differing substantially (Irwin & Sanders, Measuring Index Investment in Commodity Futures Markets, 2013). Moreover, the implied positions obtained in such way differ substantially from IID data. The problem of the Masters implied positions that a small deviation of weight of agricultural commodities (weight of feeder cattle is less than 2%) leads to a large deviation of estimation of the crude oil index trader positions. Furthermore, the Masters procedure is even rougher for an estimation of net index trader positions. To manage these problems I propose to utilize the monthly IID data since 2010

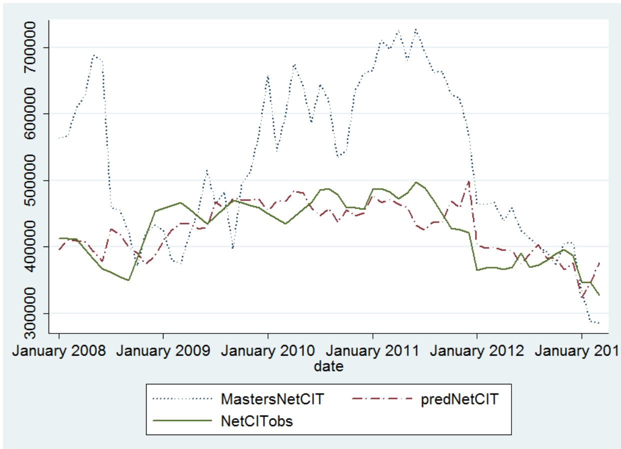
[24] Comprehensive Review of the Commitments of Traders Reporting Program http://www.cftc.gov/ucm/groups/public/@commitmentsoftraders/documents/file/noticeonsupplementalcotrept.pdf

[25] According to the explanatory notes of the CFTC, these data also have some limitations. It is possible that there are traders that replicate indices, but have not yet been identified by the CFTC staff and did not get the "special call". In addition, small traders also can implement index trading strategy, but they are not included in the IID reports. Ultimately, it is possible that entities can misunderstand what information the "special call" required them to produce.

[26] The monthly weights of each commodity in the GSCI are present in Bloomberg since 2007. The weights of commodities in the DJ-UBSCI are obtained by a request from "ETF Securities" (data is also since 2007).

and predict IID positions for the period 2007-2010 by using the Masters implied positions and quarterly observations.

Figure 7. Graph of predicted net IID positions (predNetCIT), net implied Masters positions (MastersNetCIT) and net observed IID positions (NetCITobs).



I suggest to look into the procedure of IID prediction for crude oil as an example. First, I calculate the Master implied estimation of long and short index trader positions of crude oil. Then I join linearly quarterly observations before June of 2010. I estimate regression of IID long positions against implied Masters positions obtained, time, and multiplication of time and implied Masters positions (assuming linear time-varying changes of marginal effects). At this moment, I can predict IID long positions by using the results of the regression. After this step, I estimate a regression again, but for IID short positions. In addition to implied Masters positions, I include in the regression the predicted long IID positions from the previous regression. Accordingly, I can predict short IID positions. By subtracting it from the predicted long IID positions I obtain predicted net IID positions (see all OLS estimations for four energy commodities and copper in Appendix, Table 7). Figure 5 shows the predicted net IID in a comparison with implied and observed data.

In this research monthly data is analyzed. The data about stocks of several commodities and IID positions is absent on a weekly level. Finally, I included the summary statistics of the share of net index trader position in a total open interest. The statistics of this variable for each commodity is present in Appendix (Table 5). The copper market has the highest average share of (net) index trader positions. At the same time, the cocoa market attracts the lowest number of contracts of index traders relative to the total amount of open contracts. The net index trader share ranges over 30%-40% of total open interest for the most commodity markets. Noteworthy, the cocoa and coffee markets are the smallest among considered markets. The largest market, the crude oil market is three times larger the second market, the soybeans market. The corn market has grown dramatically, because of a rise of the demand on ethanol.

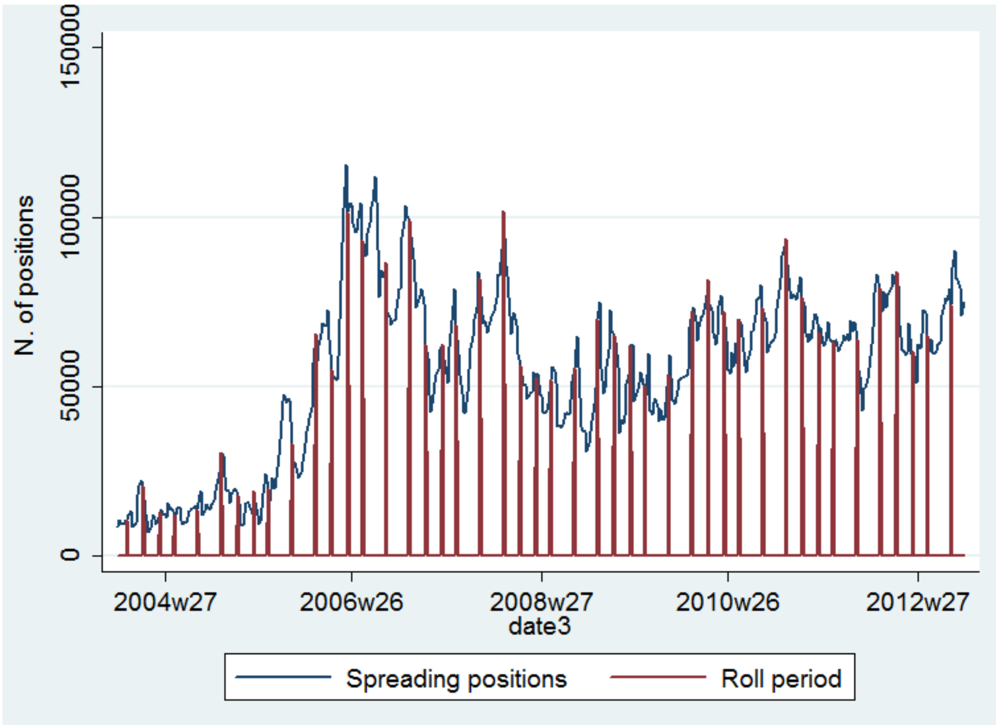
Spreading positions

The weekly data reported each Thursday is available in Thompson Reuters Datastream. The COT report which disaggregates all large traders on commercials and non-commercials contains the information about the spreading positions of non-commercial traders. The spreading positions are positions in the opposite directions on equal number of contracts. In the current literature the COT reported data is used as a proxy for positions implemented the front-running Goldman roll strategy (Frenk & Turbeville, 2011; Mou, 2011). As Figure 5 shows, the number of spreading positions has grown up in a recent time. At the same time, the character of changes of these positions has been changed. Figure 8 is typical for all agricultural commodities and indicates that the number of spreading position of wheat rises before or during the Goldman roll and then falls. For energy commodities the rolling is carried out every month, so an increase of positions before the Goldman roll concurs with a decrease of positions after the Goldman roll. Changes of spreading positions before 2004 do not have a similar pattern (Appendix, Figure 4). These graphs also imply that speculators do not try to push the spread to the fundamental value. If they push the spread to the equilibrium value, the number of spreading positions should grow during the rolling or after the rolling and not before the rolling. It is observed that they open spreading positions before the rolling.

The COT reported spreading positions are the aggregated positions of large non-commercial traders. Obviously, traders are not homogeneous and use a variety of strategies. The disaggregated COT report breaks up traders taking such positions of non-commercials managed money traders and other traders and discloses spreading positions of swap dealers. Frenk and Turbeville (Frenk & Turbeville, 2011) postulate that swap dealers that have the OTC positions on GSCI and DJ-UBSCI have a superior market information advantage. They know directly client-specific information about the size of the index trader swaps. Furthermore, all profits and losses of the roll trades go to the account of the commodity index traders and not to the swap dealers. The authors deem that swap dealers can be involved in the front-running strategy. However, the number of their spreading positions is small in a comparison with that of non-commercial traders (Appendix, Table 11). As a result, a sum of swap dealer and non-commercial positions is used as a proxy variable for traders implemented the front-running Goldman roll strategy.

The summary statistics about the spreading position share to the open interest is present in Appendix, Table 11. The share of spreading positions is lower than the share of index traders to the open interest for all commodities except natural gas and crude oil. For agricultural commodities the share of spreading positions fluctuates between 10% and 20%. Spreading positions of crude oil and natural gas account for more than 40% of their open interest. The correlation of spreading position share is negative for a majority of commodities. Because both spreading positions and index trader positions are parts of the open interest, an increase of one reduces the share of another. The correlation of spreading positions with index trader long positions is mainly positive for a majority of commodities. At the same time, for several commodities the correlation between positions is negative. Speculators may implement the front-running Goldman roll strategy with a different extent irrespectively on index trader positions.

Figure 8. The spreading positions of wheat, 2005 – 2013, weekly, Datastream, CFTC.



Methodology

In examining the impact of one variable to another, the problem of stationarity of variables is always intricate in time series regressions. Neglecting the non-stationarity of variables in times series regressions may lead to the spurious results (Granger & Newbold, 1974). Before testing the explanatory power of the fundamental factors and index trader positions on the spread, I checked their stationarity by using Augmented Dickey Fuller test. The macroeconomic factors, yield spread and real interest rate, are not stationary. The procedure and results of the estimation are described in Appendix (p. 40). The graph of the adjusted interest rate and the yield spread reflects considerable changes of variation and mean (Appendix, Figure 6). The non-stationarity of the yield spread and interest rate has been observed in the earlier literature (Bowsher & Meeks, 2006). In addition, the share of net index trader positions is also non stationary for several commodities (see Table 5 in Appendix). The stocks, hedging pressure are also non-stationary for several commodities even for the short period (2007- 2013).

To manage non-stationarity some researchers use a first difference of variables instead of level, which is usually stationary. In my case, all variables are stationary in first differences for all time periods and for all commodities. Yet, this approach does not investigate an impact of levels of variables. It is important for my case, because net index trader positions may influence levels. Moreover, I carried out the preliminary analysis and examined the OLS regressions with a first difference. The residuals estimated in such regressions have a strong autocorrelation. To get rid of autocorrelation, the lagged variables should be included in the model. Another approach is to detrend non-stationary variables. However, it is possible to do only with stationary variables with a trend. Graphs of yield spread, interest rate and cotton stocks confirm that these variables do not have a trend even for a short period 2007-2013 (Appendix).

The approach that allows working with non-stationary variables has been suggested by Pesaran and Shin (Pesaran & Shin, An Autoregressive Distributed Lag Modelling Approach to Cointegration Analysis, 1999), which is called as the autoregressive model distributed lag model (ARDL). The ARDL model is applicable for testing the causal relationship irrespective of whether the underlying regressors are purely I(1), purely I(0) [27] or mutually cointegrated[28]. However, it requires that the variables do not have the order of integration higher than one. I checked that all considered variables are at least I(1). In a comparison with models of Johansen and Engle-Granger, the ARDL does not require all regressors to be of the same order of integration. It is a great advantage for my variables, because some of them are I(1) and others are I(0). Pesaran and Shin show by using Monte Carlo experiments that the ARDL outperform the modified OLS approach of Phillips and Hansen for the small sample. The authors also argue that the ARDL corrects the problems of endogeneity of regressors. Later, Pesaran et al (Pesaran, Shin, & Smith, Bounds Testing Approaches to the Analysis of Level Relationships, 2001) have developed the ARDL model and have introduced the bound testing approach which eases the implementation and understanding of the model. For instance, when Y is a dependent variable, X and Z are regressors (scalars), the model has the following form:

$$\Delta Y_t = c_0 + c_1 t + \delta_1 Y_{t-1} + \delta_2 X_{t-1} + \delta_3 Z_{t-1} + \sum_{i=1}^{p} \theta_{1i} \Delta Y_{t-i} + \sum_{i=1}^{q} \theta_{2i} \Delta X_{t-i} + \sum_{i=1}^{r} \theta_{3i} \Delta Z_{t-i} + \psi D_t + u_t \qquad (3)$$

[27] I(0) is a stationary process, I(1) is a non-stationary process, but the first difference of the process is stationary.
[28] The processes are mutual countegrated if there is a linear combination of processes which is stationary.

D is a vector of exogenous variables such as the structural change dummies. The delta indicates the first difference operator. The order *p, q, r* is chosen by Schwarz Bayesian Criterion (SC) and Akaike Information Criterion (AIC) with the restriction that the residuals of this model should be stationary and not correlated. Monte Carlo simulations carried out by Pesaran address the evidence in favor of the SC. The model is estimated by the OLS. Pesaran and Shin point out that the coefficients are asymptotically normal. To check whether there is a level relationship, the null hypothesis of an absence of cointegration ($\delta_1= \delta_2= \delta_3=0$) is performed against a presence of cointegration ($\delta_1\neq \delta_2\neq \delta_3\neq 1$). However, the F-statistics have non-standard distribution. The critical values have been obtained by Monte Carlo simulations and are published (Pesaran, Shin, & Smith, Bounds Testing Approaches to the Analysis of Level Relationships, 2001). They depend on whether regressors are I(1) or I(0), a presence of a trend and an intercept in the model and a number of regressors. If not all regressors are I(1) and the critical value falls within the critical value for all I(0) regressors and for all I(1) regressors, the bounds test is inconclusive. If it falls higher than I(1), the null hypothesis is rejected and one can conclude about a long-run relationship between variables. If it is lower I(0), the null hypothesis cannot be rejected.

If the hypothesis of an absence of cointegration is rejected, one can evaluate the magnitude of the long-run impact of dependent variable on the independent variables. The following OLS regression is estimated:

$$Y_t=c_0+c_1t+\sum_{i=1}^{p}\alpha_i Y_{t-i}+\sum_{i=1}^{q}\beta_i X_{t-i}+\sum_{i=1}^{r}\gamma_i Z_{t-i}+\psi D_t+u_t \tag{4}$$

The parameter *p, q,* and *r* are chosen again according to the SC and the AIC. The residuals should not be autocorrelated. Then, one can calculate the magnitude of the long run effect. For example, the impact of X on Y:

$$\beta =\frac{\sum_{i=1}^{p}\beta_i}{1-\sum_{i=1}^{p}\alpha_i} \tag{5}$$

The coefficient β reflects the long-run response of Y to a unit change of X. To assess the short-run effects, the OLS regression of the error correction model is estimated:

$$\Delta Y_t=c_0+c_1t+\delta E_{t-1}+\sum_{i=1}^{p}\theta_{1i}\Delta Y_{t-i}+\sum_{i=1}^{q}\theta_{2i}\Delta X_{t-i}+\sum_{i=1}^{r}\theta_{3i}\Delta Z_{t-i}+\psi D_t+u_t \tag{6}$$

E is estimated residuals of the model (4). The coefficient δ represents the speed of convergence to the long-run equilibrium. The coefficients of other variables represent a short-run impact on the dependent variable.

To test the first hypothesis and check whether there is an impact of index traders on the spread, the following model is estimated by the OLS:

$$\Delta Spread_t = \delta_1 Spread_{t-1} + \delta_2 NetCITshare_{t-1} + \delta' Controls_{t-1} + \sum_{i=1}^{p} \theta_{1i} \Delta Spread_{t-i} + \sum_{i=1}^{q} \theta_{2i} \Delta NetCITshare_{t-i} + \sum_{i=1}^{r} \theta' \Delta Controls_{t-i} + \psi D_t + u_t \qquad (7)$$

"*Controls*" is the matrix, which columns are fundamental factors: hedging pressure, open interest, stocks, adjusted Treasury bill and yield spread. The variable "*Spreadingshare*" replaces "*NetCITshare*" to examine the second hypothesis. To test the third hypothesis both these variables are included.

Empirical results

Implementation

Before estimating the results of the regression (7), I analyze an impact of control variables on the spread in regression with a long period sample. The anticipated sign of the marginal effect of the factors is confirmed for the majority of commodities. At the same time, the coefficients are not significant for several commodities or have the opposite sign. This may be associated with the fact that the lagged spread is included in the regression and its impact is captured by the lagged spread.

To choose the order of the first differences, I apply the SC and AIC for the ARDL model (7). However, in all regressions the inclusion of an extra lag for control variables and net index trader share raises the SC and AIC values. As a result, I consider the model (3) with p equal to one, and q, r, l, m and n equal to zero as a base model. I exclude the intercept from the model. The intercept in the model (3) reflects a time trend of the spread, which should be explained by independent variables if present. In a preliminary analysis, I have included the intercept, but it is insignificant in almost all cases.

For all models, I have checked the absence of autocorrelation by the Durbin-Watson criterion. I have also checked it by using the graph of autocorrelation with the confidence interval. The hypothesis of non-stationarity has been carried out by the ADF test. In all regressions, the hypothesis of non-stationarity and the presence of autocorrelation of residuals are rejected at 5% level. In addition, I verify the normality of residuals by using the Kolmogorov-Smirnov test and the histogram of their distribution. If I observe the extreme observations, I remove these observations from the sample[29]. For several commodities, I observe that it is impossible to get rid of heteroskedasticity of residuals. As a result, I applied the robust variance estimation in Stata, which diminishes the effect of heteroskedasticity on the t-statistics.

[29] I removed 3 observations for Cocoa (4/30/08, 2/27/09, and 11/30/09), 1 observation for natural gas (8/12/09).

To evaluate the quantitative impact of net index trader share, I need to check whether there is a long-run cointegration. I compare the obtained F-test value with critical values of Pesaran (Pesaran, Shin, & Smith, Bounds Testing Approaches to the Analysis of Level Relationships, 2001). Then, I choose the order of the variables by the SC and AIC criteria and estimate the OLS regressions (4). Finally, I calculate the marginal effect by using the formula (5). Noteworthy, the estimated coefficients in (4) should be statistically significant.

For each commodity, I include monthly dummy variables to capture seasonality effects. Then, I use the F-test to verify whether these variables are jointly equal to zero. Four commodities, gasoline, lean hogs, live cattle and cotton, result in a strong seasonality of the spread. These commodities are analyzed separately in the paper.

All commodities have a specific time of expiration of futures contracts (details see in Appendix). Furthermore, there is a different frequency of expiration of contracts. As a result, all commodities analyzed are divided into three groups: commodities that rolled every month, commodities that rolled less frequently than monthly and commodities that have strong seasonality effects. The first part consists of four energy commodities and copper. The second part includes only agricultural commodities. The commodities with strong seasonality, lean hogs, live cattle, gasoline and cotton, constitute the third part.

For commodities rolled less frequently than monthly, a multiplication of roll period dummy and net index trader share captures the effect of net index traders during the roll period. The preliminary analysis shows that net index trader share is not significant or that it even has a positive sign during the non-rolling period, because of inflating nearest to maturity futures during the roll period. This contract is next nearest to maturity during the rolling period.

Impact of net index trader positions during the Goldman roll

The empirical tests support the first hypothesis. Index traders have a negative impact on the spread. The results of the estimation of the model (7) are included in Appendix (Tables 14-16). Table 14 of Appendix displays the results of estimation of the OLS for the commodities rolled every month. Table 15 of Appendix shows the results for the agricultural commodities and Table 16 of Appendix does for the commodities with a strong seasonality. In all tables, the results of the base model are displayed in the left column. I use the F-statistics to exclude variables that are not significant. The estimations of the obtained models are shown in the right column for each commodity. Table 2 shows the results of the estimated coefficients and t-statistics for the first difference and lagged level variable of the share of net index trader positions.

The 8th business day has been chosen to test whether index trader positions influence the spread during the rolling period. The 8th business day falls in with the middle of the period of rolling the GSCI and DJ-UBSCI. However, the open interest and index trader position data is published only on Thursdays of each week. The data regarding stocks, inflation, and yield spread is available only monthly. Furthermore, the estimated prediction of the IID positions of energy commodities and copperis forecasted to the end of month. As a result, the data from the end of the month or from the last Thursday is used for the analysis.

Table 2. The impact of the share of index trader positions on the spread during the rolling period.

The results of the OLS estimation of the model (7) for the share of net index trader positions are shown in this table. The dependent variable is the first difference of the spread. These estimations are obtained after removing insignificant control variables (the second column in Appendix, Tables 14-16). The monthly observations on the 8th business day of the month are considered. The independent variables are the first difference of net index trader positions share (dNetCITshare) and the lagged share of net index trader positions (L.NetCITshare). If the rolling happens less frequently than every month, the multiplication of the share of net index trader position and the roll period dummy is used. The other control variables are suppressed and present in Appendix, Tables 14-16. First number in each cell is the OLS estimation of the coefficient for the corresponded variable. A number in brackets is the t-statistics. "N" is the number of observations. The results of the test of absence of cointegration of the lagged level variables are included in the "Cointegration" row. "Yes" means that the null hypothesis about of absence of cointegration is rejected. "No" means that the null hypothesis can't be rejected or the result of the test is inconclusive (Pesaran & Shin, An Autoregressive Distributed Lag Modelling Approach to Cointegration Analysis, 1999). The symbols *, **, and *** denote statistical significance at the 10%, 5% and 1% levels.

	Heat. Oil	Nat. Gas	Crude Oil	Copper	Cocoa	Coffee	Corn
dNetCITshare	0.010	-0.515**	0.017	-0.002	-0.030*	-0.002	0.021
	(0.28)	(-2.49)	(0.40)	(-0.71)	(-1.74)	(-0.26)	(0.53)
L.NetCITshare	-0.070***	-0.024	-0.011**	-0.003**	-0.003	-0.002***	-0.004
	(-3.66)	(-1.08)	(-2.39)	(-2.11)	(-1.30)	(-3.42)	(-0.89)
N	70	68	70	70	87	87	83
adj. R^2	0.383	0.508	0.198	0.523	0.289	0.430	0.566
Cointegration	Yes***	Yes***	No	No	Yes*	Yes***	Yes***

	Gasoline	Lean Hogs	Live Cattle	Cotton	Soybeans	Sugar	Wheat
dNetCITshare	0.010	-0.038	-0.093**	-0.066	-0.000	-0.124**	-0.018
	(0.18)	(-0.59)	(-2.01)	(-0.96)	(-0.03)	(-2.01)	(-0.90)
L.NetCITshare	-0.099**	-0.068*	-0.067***	-0.030	-0.006*	-0.012	-0.006***
	(-2.72)	(-1.85)	(-3.17)	(-1.05)	(-1.82)	(-1.62)	(-2.79)
N	70	85	87	87	86	86	87
adj. R^2	0.903	0.833	0.772	0.228	0.163	0.173	0.223
Cointegration	Yes***	Yes***	Yes***	Yes***	Yes*	No	Yes**

The main variable analyzed is the lagged net index trader share. This variable represents the impact of changes of level of net index trader positions on the spread. The increment of net index trader share has a secondary meaning. When the index trader positions increases in a period between rolling periods, it pushes the price of nearest to maturity futures up. It may result in a more backwardated spread.

The impact of net index trader positions share is significantly negative on the spread in the base model for coffee, wheat, heating oil, gasoline and live cattle markets. For all commodities except cocoa, cotton, corn, sugar and natural gas, the negative impact is significant at least at a 10% level of confidence in the model with excluded insignificant control variables. It is worth saying that for cocoa, sugar and natural gas, the first difference of net index trader share is negatively significant at the 10% level of confidence. The impact of net index trader share is negative for all commodities. The negative impact on the spread means that higher share of index traders induces a commodity market to be less backwardated. This result is coincided with the results of Brunetti and Reiffen (Brunetti & Reiffen, 2011). Index traders take a short position for the nearest-to-maturity futures contract and push down the price of the contract. At the same time, they take a long position for the first deferred futures contract and push up the price. As a result, the spread became more negative.

The coefficients have the following economic interpretation. The coefficient of the share of net index trader positions divided by the coefficient of lagged spread (it is present in Appendix) and multiplied by a negative sign gives a rough estimation of the long-run marginal impact. However, these results are valid only if the hypothesis of absence of cointegration is rejected in the model (7). Moreover, it is a rougher estimation compared with the model (5). The coefficient of the first difference reflects the short-run effect. This effect represents the impact of changes of index trader positions in the last month on changes of the spread in the current month. These coefficients are mainly not significant, because the changes of the share of index trader positions are relatively small for a majority of commodities compared with the mean of the share of these positions (Table 10 of Appendix). The short-run impact might be captured by fundamental or seasonality factors. It is worth saying that the standard deviation of the share of net index trader position of natural gas and sugar are the highest (10% and 5.9%, Table 10 of Appendix). Both of these commodities have significant and comparable results. An increase of the share of index trader positions by 5% leads to a decrease of spread by 2.5% for natural gas and 0.6% for sugar in the current month if other factors are unchanged. These results should be compared with the mean and standard deviation of the corresponding commodities (Table 8 of Appendix). Natural gas has a much lower mean (-3.2%) and higher standard deviation of the spread (4.6%) in a comparison with sugar (mean is -0.1% and volatility is 1.8%), because of higher storage costs for natural gas.

Table 3. The long-run effect of the net index trader position share on the spread.

The OLS regressions are estimated only for those commodities that have statistically significant coefficient of "NetCITshare" and reject the hypothesis about the absence of the cointegration. The dependent variable is the spread (Spread). The independent variables are lags of the spread and the share of net index trader positions. The control variables are suppressed and included in Appendix, Table 17.The estimation is carried out for the variables that are significant in the right column of each commodity in Tables 14, 15 and 16 of Appendix. The order of the lagged variables was chosen by the SC and AIC. The long-run marginal effect is calculated by the formula (5). The impact of one standard deviation of the net index trader share ("1 sd impact") is computed as a multiplication of the long-run marginal effect and the standard deviation of "NetCITshare" from Table 10 of Appendix. The standard deviation of the spread of the corresponding commodity is attached in the "Sd of spread" line. Adjustment coefficient is the estimated coefficient δ in the model (6). It represents the speed of convergence to the equilibrium. The symbols *, **, and *** denote statistical significance at the 10%, 5% and 1% levels.

	Gasoline	Heating Oil	Lean Hogs	Live Cattle	Coffee	Soybeans	Wheat
L.Spread	0.308**	0.248**	0.537***	0.716***	0.670***	0.710***	0.544***
	(2.58)	(2.24)	(5.61)	(6.73)	(8.49)	(5.78)	(3.57)
L2.Spread		0.416***		-0.240*	0.285***		0.331**
		(3.75)		(-2.27)	(3.40)		(2.27)
NetCITshare	-0.069**	-0.066***	-0.075**	-0.066***	-0.002**	-0.006***	-0.005***
	(-2.25)	(-3.84)	(-2.20)	(-3.46)	(-2.83)	(-2.93)	(-3.06)
N	70	71	86	87	87	87	87
adj. R^2	0.861	0.766	0.860	0.760	0.980	0.589	0.885
Long-run marginal effect	-0.099	-0.196	-0.164	-0.126	-0.045	-0.021	-0.045
1 sd effect	-0.473%	-0.583%	-0.838%	-0.710%	-0.269%	-0.129%	-0.255%
Sd of spread	3.483%	1.296%	3.750%	1.401%	0.425%	1.059%	0.615%
Adjustment coefficient	1.188***	1.150***	0.752***	0.765***	-0.146	0.870*	0.506*

The estimation of the model (5) is required to answer the question "what is the magnitude of the long-run impact of index traders". Not all commodities with significant results are cointegrated with the spread. The F-test has inconclusive results for crude oil and copper. Consequently, I estimated the marginal effect of net index trader position share only for six commodities with cointegrated variables. The results are shown in Table 3 (all results with control variables are in Table 17 of Appendix). To compare the estimated long-run marginal effect among commodities, the impact of one standard deviation of the share of net index trader positions is calculated. The negative impact of one standard deviation of net index trader share varies from 0.12% to 0.84% of the spread. The long-run effect has an effect on the dependent variable (Spread) over future time periods at a rate of the convergence to equilibrium (Adjustment coefficient) per time period. It is observed in Table 3, that energy commodities have a higher coefficient of adjustment, which is almost equal to one. This means that the impact of one standard deviation is persistent over time and has the same magnitude. Agricultural commodities have much lower, or not even significant coefficients of adjustment. It may be connected with the fact that they are rolled less frequently than every month. Moreover, livestock commodities are rolling more frequently than grains. As a result, their coefficient is higher. This result suggests that the impact of index traders decays or even disappears in the non-rolling periods.

To clarify how to calculate the impact of index traders over time, lean hogs is considered as an example. The adjustment coefficient of lean hogs is 0.752 and long-run marginal effect is -0.164. This means that, in the current month, an increase of the share of index traders by 1% leads to a decrease of the spread by 0.12%. In the next month, it leads to a decrease of the spread by 0.9%. After five months, it decreases by 0.04%. However, the decaying impact of index trader positions is likely to be explained by the frequency of rolling per year. To adjust for the frequency of rolling, one can assume that the convergence for equilibrium is near one, which is supported by the energy commodities. These commodities are rolled every month.

To understand the economic significance, this impact should be compared with the standard deviation of the spread. The results are varied among commodities. The effect of the net index trader share is stronger for less liquid[30] commodities such as lean hogs and live cattle. The effect is the lowest for the most liquid commodity, soybeans. Gasoline and heating oil, which are also relatively liquid, have a higher impact compared with soybeans, but they have a higher volatility of the spread. The commodities that do not have significant and cointegrated results are also more liquid than live cattle and lean hogs. At the same time, the proper test for this hypothesis would be a cross-sectional regression with factors capturing liquidity. The sample of commodities analyzed is not large enough to carry out such regression analysis.

The losses due to an increase of index trader share are large enough for passive investors. For heating oil, the losses due to one monthly standard deviation make up to 0.58% (or even higher if consider the coefficient of adjustment higher than one) per month. Compared with the average changes of nearest-to-maturity futures price per month equal to 0.8% for heating oil (Table 3), the losses due to the share of index traders have a great impact. Lean hogs and live cattle have even higher losses than monthly futures returns over the period 2006-2013.

Table 4. The average monthly futures return of commodities over the period 2006- 2013

	Gasoline	Heating oil	Lean Hogs	Live Cattle	Coffee	Soybeans	Wheat
Futures return[31]	0.64%	0.81%	0.31%	0.31%	0.04%	0.99%	0.79%

[30] The liquidity of the market is defined by the dollar value of open interest. The estimations for 2012 are present in Appendix, Table 10.
[31] The average monthly futures return (changes of nearest-to-maturity futures price) over the period 2006-2013.

This result implies that investing in recently introduced indices which roll futures in long-maturity futures does not solve the problem of negative roll returns. First, a relatively small amount of inflow investing in such indices may compose a great share of open interest of long-maturity indices. As a result, the negative impact of index traders can be even higher compared with the traditional indices. Second, long-maturity futures are much less liquid. If my observation about liquidity effects can be generalized among all commodities, the negative impact of rolling might be even higher.

The fact that the results are not significant for several commodities may be caused by activity of speculators, which is discussed below. Also, several articles theoretically show that the impact of speculators can be higher for the period of supply and demand disruptions (Basak & Pavlova, A Model of Financialization of Commodities, 2013; Brunetti & Reiffen, 2011). At such periods, signals of fundamentals coincide with the period of rolling of index traders. The impact of index traders is misinterpreted with fundamentals by traders and distorts the prices. As a result, decreases of net index trader share may lead even to higher impact in a comparison with periods without supply and demand disruptions. Controlling for supply and demand disruptions is an intricate task. The level of inventories has conceptual lacks and timing issue (Gorton, Hayashi, & Rouwenhorst, The Fundamentals of Commodity Futures Returns, 2007). Information about the level of inventories is often published with a time lag and subsequently revised.

The impact of other variables is frequently not significant. As I noted above, it may be caused by the inclusion of the lagged dependent variable. The lagged dependent variable captures their impact. It is worth saying that signs of coefficients of control variables are mainly corresponded with the signs observed in the current literature (Tables 14-16 of Appendix). Especially, it concerns the short-run impact. For example, the first difference of stocks is significantly negative for six commodities, while those of hedging pressure are significantly positive for four commodities. At the same time, the long-run impact of control variables is significant for a few commodities and it does not always have the anticipated sign of correlation with the spread. It is worth noting that the contemporaneous correlation of the control variables and the spread or the futures return has been investigated in the current literature. The long-run impact of control variables is not properly covered in the literature and is not easily interpreted.

Impact of spreading positions before the Goldman roll

The results provide evidence in favor of the second hypothesis. Speculators have a negative impact on the spread before the Goldman roll. The positions on the first business day of the month when the rolling happens are investigated. Table 5 shows the main results of Tables 18-20 of Appendix. Table 18 of Appendix reports the OLS estimation of the commodities rolled every month. Table 19 of Appendix displays the results for the agricultural commodities, while Table 20 of Appendix shows the estimations for the commodities with a strong seasonality. The main variable is the spreading positions share.

As noted above, a sum of spreading positions of non-commercials and swap dealers is assumed to be a proxy for spreading positions implementing the front-running Goldman roll strategy. It is a proxy for two reasons. First, it is an aggregated data of all spreading positions. Therefore, many positions included are associated with the front-running strategy. Second, there is no data about intraday spreading positions that also can implement the front-running strategy. It is assumed that they are positively correlated with the total number of spreading positions.

Table 5. The impact of the share of spreading positions on the spread before the Goldman roll.

The results of the OLS estimation of the model (7) for the share of spreading positions are shown in this table. These estimations are obtained after removing insignificant control variables (the second column in Tables 18-20). The monthly observations on the 1st business day of the month are considered. The dependent variable is the first difference of the spread. The independent variables are the first difference of the share of spreading positions (dSpreadingshare) and the lagged share of spreading positions (L.Spreadingshare). If the rolling happens less frequently than every month, the multiplication of the share of spreading positions and the roll period dummy is used. The other control variables are suppressed and present in Tables 18-20 of Appendix. First number in each cell is the OLS estimation of the coefficient for the corresponded variable. A number in brackets is the t-statistics. "N" is the number of observations. The results of the test of the absence of cointegration of the lagged level variables are included in the "Cointegration" row. "Yes" means that the null hypothesis about of absence of cointegration is rejected. "No" means that the null hypothesis can't be rejected or the result of the test is inconclusive (Pesaran & Shin, An Autoregressive Distributed Lag Modelling Approach to Cointegration Analysis, 1999). The symbols *, **, and *** denote statistical significance at the 10%, 5% and 1% levels.

	Heat. Oil	Nat. Gas	Crude Oil	Copper	Cocoa	Coffee	Corn
dSpreadingshare	-0.066**	0.063	0.086***	-0.014**	-0.006	-0.011	-0.091**
	(-2.86)	(0.83)	(3.50)	(-2.35)	(-0.38)	(-1.09)	(-2.57)
L.Spreadingshare	-0.052***	-0.015*	-0.040***	0.006*	-0.016***	-0.006***	-0.009
	(-5.10)	(-1.80)	(-3.65)	(1.81)	(-3.93)	(-2.91)	(-1.30)
N	74	69	74	75	82	82	77
adj. R^2	0.410	0.363	0.445	0.454	0.416	0.382	0.550
Cointegration	Yes***	Yes***	Yes***	Yes*	Yes***	Yes***	Yes***

	Gasoline	Lean Hogs	Live Cattle	Cotton	Soybeans	Sugar	Wheat
dSpreadingshare	0.028	-0.077	0.018	-0.118*	0.009	-0.025	0.031
	(0.46)	(-0.60)	(0.29)	(-1.94)	(0.30)	(-0.23)	(1.25)
L.Spreadingshare	-0.050	-0.060	0.013	0.024	-0.022***	-0.033*	-0.026***
	(-0.78)	(-0.78)	(0.26)	(0.41)	(-3.49)	(-1.81)	(-4.50)
N	74	80	82	82	82	81	82
adj. R^2	0.868	0.844	0.674	0.343	0.174	0.174	0.426
Cointegration	Yes***	Yes***	Yes***	No	Yes**	Yes*	Yes***

The spreading positions share has a negative significant impact for heating oil, cocoa, coffee, corn, soybeans and wheat at least at a 10% confidence level in a base model. After excluding insignificant controls, natural gas, crude oil and sugar show evidence of a negative and significant impact. The higher share of spreading positions leads to higher contango before the rolling period. This result supports observations of Frenk and Turbeville (Frenk & Turbeville, 2011). By implementing the strategy the speculators push the price of nearest-to-maturity contracts down and the price of next-to-maturity contracts up before the rolling period.

Table 6. The long-run effect of the spreading position share on the spread.

The OLS regressions are estimated only for those commodities that have statistically significant coefficient of ”Spreadingshare” and reject the hypothesis about the absence of the cointegration. The estimation is carried out for the variables that are significant in the right column of each commodity in Tables 18, 19 and 20. The dependent variable is the spread (Spread). The independent variables are lags of the spread and the share of spreading positions. The control variables are suppressed and included in Appendix, Table 17. The order of the lagged variables is chosen by the SC and AIC. The long-run marginal effect is calculated by the formula (5). The impact of one standard deviation of the net index trader share (“1 sd impact”) is computed as a multiplication of the marginal effect and the standard deviation of ”Spreadingshare” from Table 11 of Appendix. The standard deviation of the spread of the corresponding commodity is attached in the bottom line. The t-statistics of monthly dummy variables is suppressed. Adjustment coefficient is the estimated coefficient δ in the model (6). It represents the speed of convergence to the equilibrium. The symbols *, **, and *** denote statistical significance at the 10%, 5% and 1% levels.

	Heat.Oil	Nat. Gas	Crude Oil	Cocoa	Coffee	Soybeans	Sugar	Wheat
L.Spread	0.319***	1.001***	0.819***	0.405***	0.712***	0.619***	0.779***	0.493***
	(3.30)	(7.79)	(12.87)	(4.45)	(8.37)	(5.11)	(8.73)	(3.04)
L2.Spread		-0.399***		0.248**				
		(-7.02)		(2.62)				
Spreadingshare	-0.045***	-0.017***	-0.020*	-0.016***	-0.006***	-0.018***	-0.039**	-0.025***
	(-4.63)	(-2.66)	(-1.68)	(-4.36)	(-3.78)	(-3.53)	(-2.59)	(-3.55)
N	75	68	77	83	83	82	82	82
adj. R^2	0.815	0.823	0.781	0.627	0.965	0.488	0.645	0.764
Long-run marginal effect	-0.066	-0.043	-0.110	-0.046	-0.021	-0.047	-0.176	-0.049
1 sd effect	-0.240%	-0.177%	-0.405%	-0.170%	-0.072%	-0.121%	-0.627%	-0.135%
Sd of spread	1.296%	4.599%	1.739%	0.620%	0.425%	1.059%	1.841%	0.615%
Adjustment coefficient	0.995***	1.021***	1.695***	0.126	0.026	0.029	0.314	0.522*

Table 6 shows the long-run marginal effect of the share of spreading positions for significant and cointegrated results (results with all control variables in Table 21 of Appendix). The impact of one standard deviation of spreading positions share is lower than that of index trader share for all commodities. For example, for heating oil, soybeans, and wheat, the impact of speculators is more than two times lower than that of index traders during the Goldman roll. It can be explained that all commodities have a much lower share of spreading positions to the open interest than that of index trader positions. The adjustment coefficient for energy commodities is much higher than for agricultural commodities as in the previous test. Heating oil and natural gas have the adjustment coefficient close to unity. However, the coefficient of crude oil is higher than unity. It means that the impact of speculators is extending over time. The changes of structure of spreading positions can explain the phenomenon. The share of speculators implementing the front-running strategy to the number of spreading positions may be increased. As a result, the impact of one standard deviation of the share of spreading positions may go up over time.

Only five commodities do not have any evidence of a significant negative impact of speculators. Copper and cotton have a not significant level impact, but a significantly negative impact of the first difference of spreading positions share. The reason for not having significant results might be related with the choice of the day. However, I also investigated the 4th business day and found that results are still not significant for these commodities. As noted above, an aggregation of spreading positions can be a reason for not significant results.

Impact of spreading positions during the Goldman roll

It is found that speculators do not influence the spread during the Goldman roll. The results of inclusion of spreading positions on the spread during the rolling period are shown in Table 7 (the results with control variables in Tables 22-24 of Appendix). The 8th business day is considered. Table 22 shows of Appendix the OLS estimation for the energy commodities and copper. The results of estimations for the agricultural commodities are present in Table 23 of Appendix. Table 24 of Appendix displays the OLS estimations for the commodities with a strong seasonality.

All regressions except a base model for crude oil have insignificant coefficients of spreading position share. As for crude oil, excluding insignificant variables by the F-test decreases the t-statistics of the coefficient of spreading position share, which is insignificant. This means that, on average, spreading positions do not lead to higher or lower contango of the spread. Despite the impact of speculators before the Goldman roll, the variation of the spread is explained by other factors.

As it is found above, speculators implementing the front-running strategy push the spread down before the Goldman roll. However, during the Goldman roll, they close their spreading positions and provide the commodity markets with liquidity during the Goldman roll. Consequently, the total impact of spreading positions over whole period is equal to zero. Moreover, it also means that higher spreading positions leads to lower profits of speculators implementing the front-running strategy. Higher impact before the Goldman roll is compensated by providing more liquidity during the rolling period. This effect conforms to the observations of Mou (Mou, 2011), who found a decrease of profits of the front-running strategy after 2006.

The results of regression do not imply that speculators never have an impact on the spread during the rolling period. They can influence positively and negatively over time, but the average impact does not differ significantly from zero. The confirmation of the first two hypotheses implies that, if the share of net index trader positions goes down and the share of spreading positions goes up at the same rolling period, speculators push the spread too low before the Goldman roll. However, during the rolling, they push the spread up to the values explained by the impact of speculators. As a result, their total impact is not significant on the spread during the Goldman roll. According to Table 12, the spreading positions and net index trader positions are positively correlated for a majority of commodities. So it means that they commove and such cases happen rarely.

Table 7. The impact of the share of spreading positions on the spread during the Goldman roll.

The results of the OLS estimation of the model (7) for the share of spreading positions are shown in this table. The dependent variable is the first difference of the spread. These estimations are obtained after removing insignificant control variables (the second column in Tables 22-24). The monthly observations on the 8^{th} business day of the month are considered. The dependent variable is the first difference of the spread. The independent variables are the first difference of the share of spreading positions (dSpreadingshare) and that of the net index trader positions (dNetCITshare) and the lagged levels of these variables (L.Spreadingshare, L.NetCITshare). If the rolling happens less frequently than every month, the multiplication of the share of spreading positions and the roll period dummy is used. The other control variables are suppressed and present in Tables 22-24 of Appendix. First number in each cell is the OLS estimation of the coefficient for the corresponded variable. A number in brackets is the t-statistics. "N" is the number of observations. The symbols *, **, and *** denote statistical significance at the 10%, 5% and 1% levels.

	Heat.Oil	Nat. Gas	Crude Oil	Copper	Cocoa	Coffee	Corn
dNetCITshare	0.026	-0.305**	0.040	-0.002	-0.022	0.000	0.026
	(0.71)	(-2.92)	(0.80)	(-0.66)	(-1.16)	(0.01)	(0.57)
dSpreadingshare	-0.015	0.047	0.171***	-0.003	0.004	0.008	-0.008
	(-0.48)	(0.49)	(3.74)	(-0.39)	(0.22)	(0.81)	(-0.11)
L.NetCITshare	-0.052*	-0.011	-0.024	-0.003*	0.007	-0.003	-0.011
	(-2.32)	(-0.64)	(-1.24)	(-2.38)	(0.72)	(-1.38)	(-0.84)
L.Spreadingshare	-0.027	-0.000	0.009	0.005	-0.014	0.001	0.017
	(-1.46)	(-0.00)	(0.73)	(0.85)	(-1.11)	(0.16)	(0.68)
N	70	67	70	70	82	82	78
adj. R^2	0.385	0.625	0.265	0.518	0.279	0.385	0.553
	Gasoline	Lean Hogs	Live Cattle	Cotton	Soybeans	Sugar	Wheat
dNetCITshare	0.003	0.010	-0.098	-0.008	0.006	-0.125*	-0.028
	(0.06)	(0.13)	(-1.51)	(-0.13)	(0.42)	(-1.96)	(-1.05)
dSpreadingshare	-0.029	0.068	0.049	0.032	-0.010	0.097	0.011
	(-0.47)	(0.73)	(0.65)	(0.33)	(-0.30)	(0.66)	(0.39)
L.NetCITshare	-0.145***	-0.098**	-0.066***	-0.036	-0.011	0.008	-0.008
	(-3.33)	(-2.24)	(-2.81)	(-1.23)	(-1.49)	(0.44)	(-1.81)
L.Spreadingshare	-0.143*	-0.010	0.041	0.047	0.013	-0.047	0.003
	(-1.67)	(-0.14)	(0.87)	(0.48)	(1.07)	(-1.14)	(0.34)
N	70	80	82	82	81	81	82
adj. R^2	0.905	0.836	0.750	0.207	0.136	0.153	0.189

In addition, there are two observations that imply that speculators do not drive the spread to the fundamental value. First, if they do that, their effect would be positively significant. Table 6 displays the absence of the positive impact. Second, Figure 8 indicates that spreading positions increases before the Goldman roll. It would be expected that to take advantage of arbitrage from index trader activity, the speculators may open spreading positions during or after the Goldman roll. The reason of such behavior is that after the Goldman roll the liquidity is low and futures expiration is close. The third argument in a favor of this statement is that after the Goldman roll the negative impact of index traders is still significant[32], that means that their impact does not disappear after the rolling.

The practical implication for passive investors is that indices that implement different rolling schedule do not solve the problem of negative returns if these indices are popular enough. If an index uses the time which is not known for speculators, the rolling process still may have a negative impact on the roll returns of passive investors. The negative impact is defined by the share of index traders.

These results have several limitations. First, daily variations of spreading positions might be high before and during the rolling period. In this paper, weekly observations of spreading positions are used, which are not always coincided with the 8th business day. The time lag can amount to four days. Second, spreading positions used in the paper are aggregated. The changes of the structure of these positions may have an impact on the spread during the rolling period. However, the fact that spreading position share is insignificant for all commodities makes this argument weak. Otherwise, the changes of the structure should be homogeneous among all commodities included which are hardly probable. Third, there can be omitted factors that influence the spread[33]. As noted above, supply and demand disruptions can have an impact on the influence of index trader on the spread (Basak & Pavlova, A Model of Financialization of Commodities, 2013). Noteworthy, if in times of shocks of supply and demand the impact of spreading positions is higher, the results make evidence that on average for all periods the impact is insignificant. Fourth, intraday spreading positions may also have an impact on prices of commodities during the Goldman roll (Bichetti & Maystre, 2012).

[32] These results are not reported in the paper.

[33] The analysis of the interaction of net index trader share and spreading positions share was also carried out, but is not reported in the paper. The multiplication of both variables was included in the regression. However, it turns out to be not significant. It means that the level of spreading positions does not influence the marginal effect of net index trader share.

Conclusion

The impact of index traders and speculators on the roll return of passive investors is under investigation in this paper. The autoregressive distributed lag model, which allows analyzing non-stationary variation of levels of index trader activity, is applied. The negative and significant impact of index traders is observed for 12 out of 14 commodities analyzed during the Goldman roll period. To capture the impact of speculators performing the front-running strategy, the data regarding spreading positions of swap-dealers and non-commercials is considered. The spreading positions are negatively and significantly correlated with the spread before the Goldman roll for 9 out of 14 commodities. Yet, it is found that the speculators do not have an impact on the spread during the Goldman roll period. This means that push the spread down before the Goldman roll and push the spread up during the Goldman roll so that their total impact is insignificant.

The main implication for passive investors is that the share of their positions to open interest decreases their roll returns and, as a result, drops their total commodity returns dramatically. One standard deviation of the share of index trader positions drops the roll return by 0.46% on average among commodities with significant results. This decrease is comparable with average futures monthly return equal to 0.56% for the same commodities. This result infers that investing in long-maturity futures or other non-liquid futures may be even more detrimental for passive investors, because a relatively smaller number of index trader positions would have a higher share to open interest for such futures and more negative roll returns.

The second implication concerns the activity of speculators. It is found that they implement the front-running strategy and do not influence the roll return during the roll period. Their spreading positions are highly and positively correlated with the long index trader positions. For passive investors, it means that a disclosure of rolling dates and higher transparency of positions of index traders do not influence the roll return.

An ideal index for passive investors should invest in liquid futures contracts or should be not popular (among index traders) enough to invest in illiquid futures. The index should also avoid rolling futures contracts during the Goldman roll and several days before and after the Goldman roll. Noteworthy, the Goldman roll period and several days before are the most liquid period of time. The problem of liquidity can be crucial even for relatively less popular indices among index traders. If such an ideal index exists, it would become popular quickly and all benefits would disappear. Ultimately, the indices investing in the most liquid commodities and in the most liquid time would have the least negative roll return.

The approach applied in the paper has several limitations. The mismatching between publication of data about positions of traders by the CFTC and the Goldman roll period is present in the analysis. The aggregation of spreading positions can be an issue. Both these problems can be solved by using the non-published Larger Trader Reporting System (LTRS) database of the CFTC. This data is used in several articles (Brunetti & Reiffen, 2011; Buyuksahin & Robe, 2011). The omission of factors, such as intraday trading that can influence the results, are hardly probable. The impact of spreading positions is not significant for all commodities. If these factors are present, the results should be significant at least for several commodities.

Further research can be developed into two directions. First, it would be interesting to dig deeply into the problem of the large roll yield and investigate the periods when the rolling period overlaps with signals of strong fundamental factors. This research might help in finding optimal rolling periods if they exist. Second, the impact of index traders on the behavior of traditional hedgers on the commodity market, producers and utilities is up for deeper investigation. These questions give a broader view on how traders influence physical commodity prices and whether there is an impact of index traders on the spread in the non-rolling periods[34].

The negative impact of index traders during the rolling period has been also empirically shown in other articles (Brunetti & Reiffen, 2011; Buyuksahin & Robe, 2011). This paper extends the current literature with the analysis of speculators' activity. Speculators do not mitigate or exacerbate the impact of index trader on the roll yield. Instead, they spread the index trader's impact over a longer period than the Goldman roll period.

[34] The impact of index traders during the non-rolling periods has been investigated but not reported. The results turn out to be mixed and frequently not significant among commodities.

Appendix

Figure 1. Weights of GSCI and DJ-UBS indices in January 2013 (Norrish & Croft, 2013)**.**

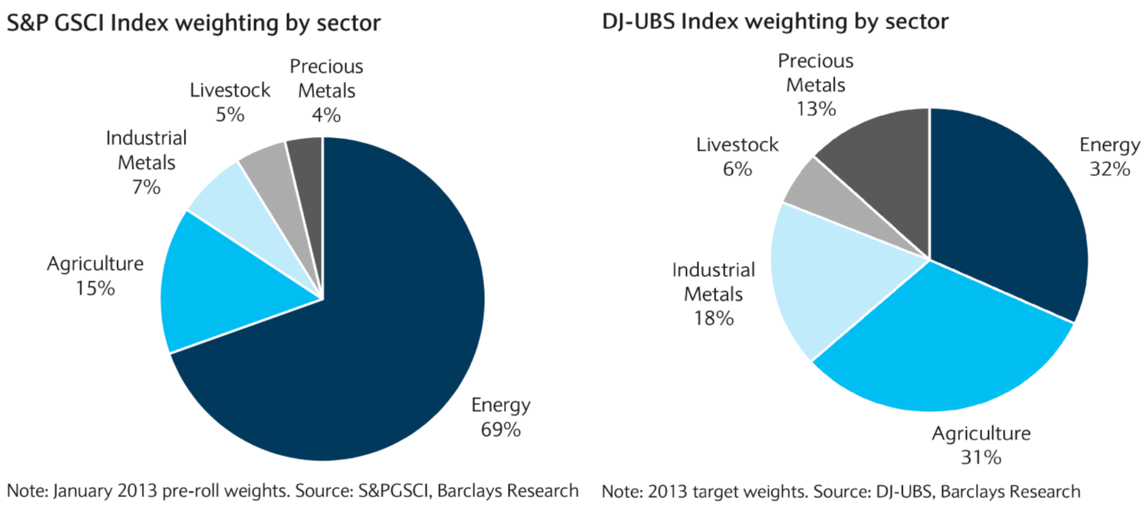


Table 1. Relationship between Legacy, Disaggregated and Supplemental Commitments of Traders Reports (Irwin & Sanders, Testing the Masters Hypothesis in Commodity Futures Markets, 2011)**.**

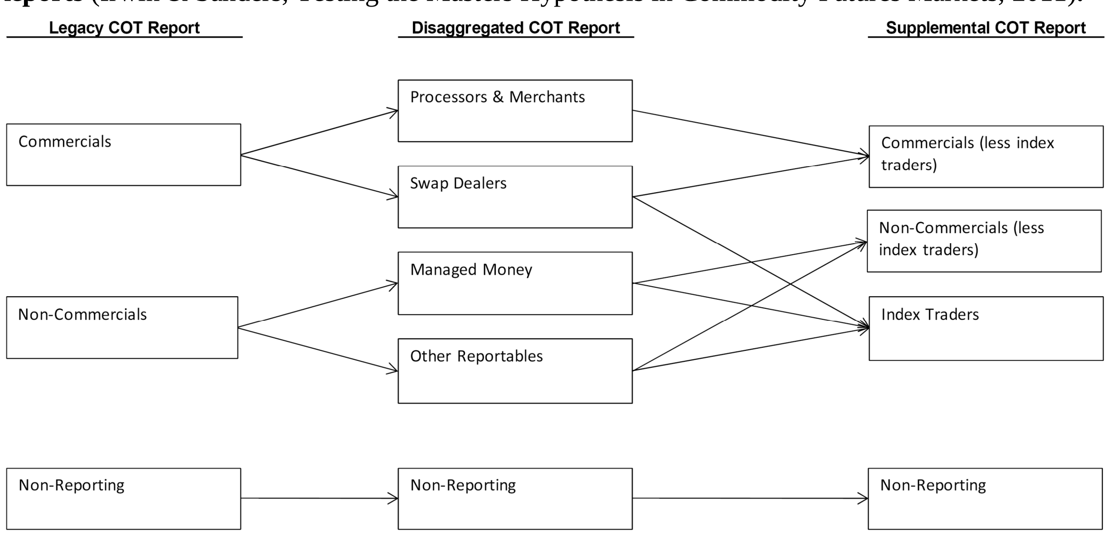


Figure 2. The ratio of spreading positions and the open interest for Lean Hogs and Cocoa, Datastream, CFTC.

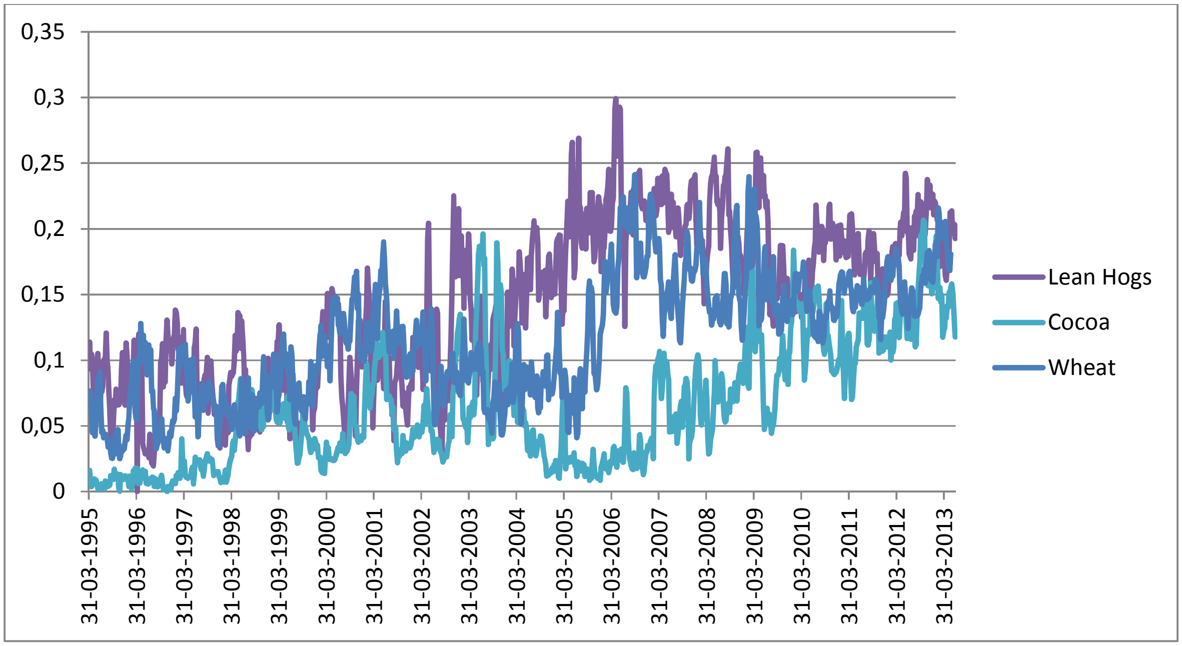


Figure 3. The ratio of spreading positions and open interest for Orange Juice and Paladium, Datastream, CFTC.

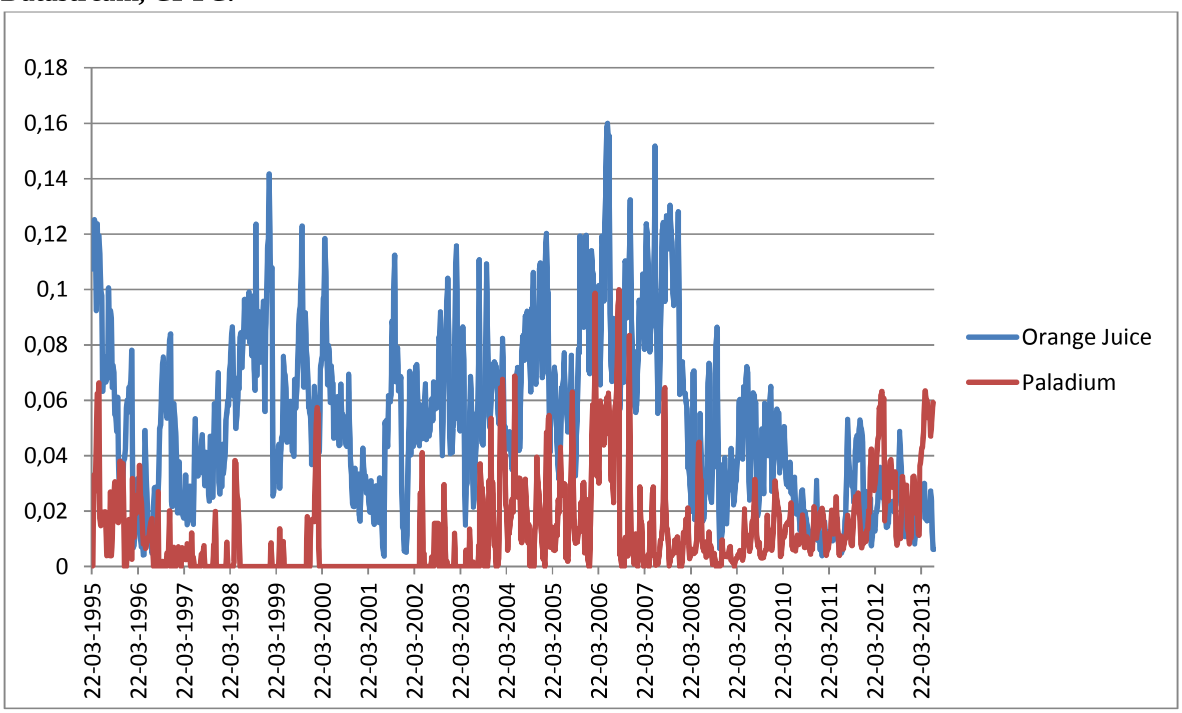


Table 2. The rolling scheme of the GSCI and the DJ-UBS indices (Irwin & Sanders, Testing the Masters Hypothesis in Commodity Futures Markets, 2011)**.**

Rolls are executed at the beginning of the corresponding months over the 5th and 9th business days for the GSCI and over 6th and 10th business days for the DJ-UBSCI. The numbers in Table designate the futures contract month in the indices as of the beginning of the month (e.g. the CBT wheat contracts are rolled from the March contract to the May contract in February each year).

Commodity	Exchange	Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug.	Sep.	Oct.	Nov.	Dec.
Wheat	CBT	3	3	5	5	7	7	9	9	12	12	12	3
Corn	CBT	3	3	5	5	7	7	9	9	12	12	12	3
Cotton #2	ICE	3	3	5	5	7	7	12	12	12	12	12	3
Soybeans	CBT	3	3	5	5	7	7	11	11	11	11	1	1
Sugar #11	ICE	3	3	5	5	7	7	10	10	10	3	3	3
Coffee "C"	ICE	3	3	5	5	7	7	9	9	12	12	12	3
Cocoa	ICE	3	3	5	5	7	7	9	9	12	12	12	3
Lean Hogs	CME	2	4	4	6	6	7	8	10	10	12	12	3
Live Cattle	CME	2	4	4	6	6	8	8	10	10	12	12	3
Crude Oil (WTI)	NYMEX	2	3	4	5	6	7	8	9	10	11	12	1
Gasoline (RBOB)	NYMEX	2	3	4	5	6	7	8	9	10	11	12	1
Heating Oil #2	NYMEX	2	3	4	5	6	7	8	9	10	11	12	1
Natural Gas	NYMEX	2	3	4	5	6	7	8	9	10	11	12	1
Copper	NYMEX	2	3	4	5	6	7	8	9	10	11	12	1

Figure 4. The average spread of agricultural commodities in the Goldman roll period and over three weeks before to two weeks after the Goldman roll before 2004.

"Average" is the average spread over the whole period before 2004[35], Datastream, weekly data.

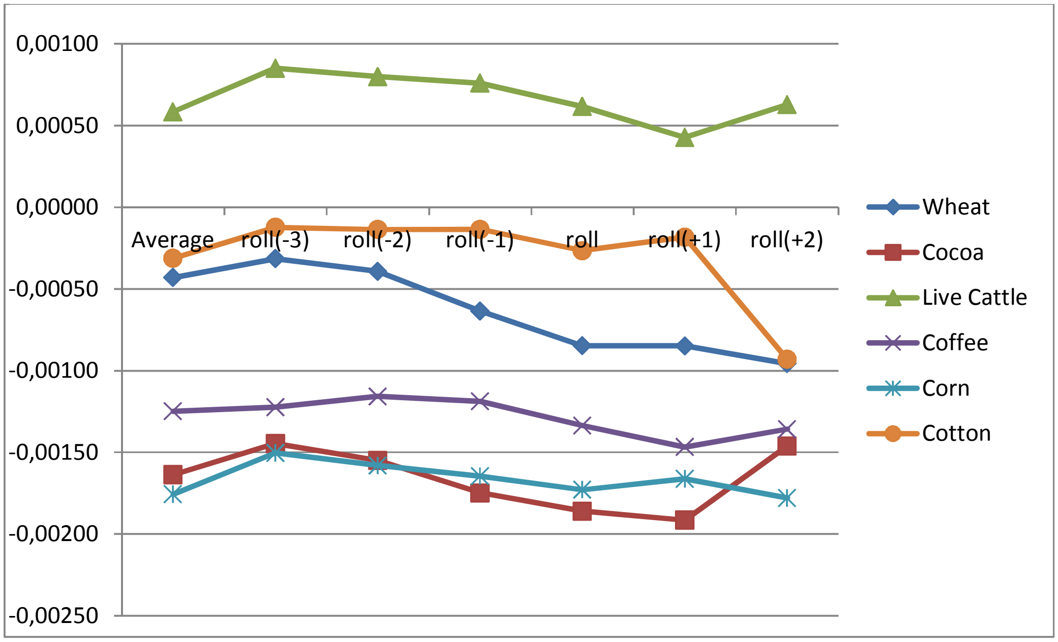
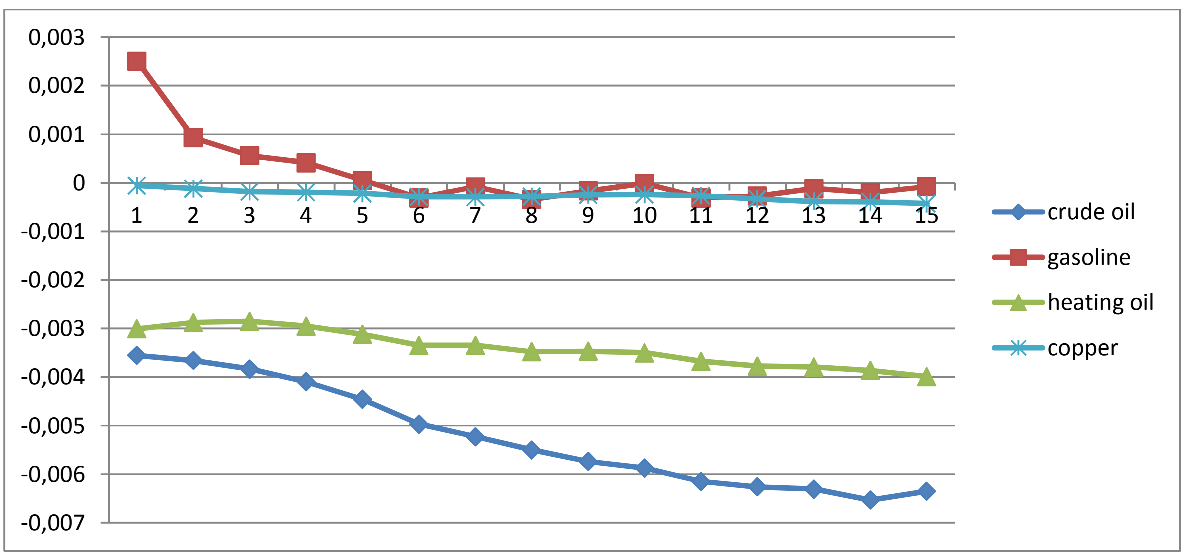


Figure 5. The average spread of energy commodities and copper in the Goldman roll period (5th - 9th business days) and over five days before to 5 days after the Goldman roll, Datastream, daily data.

The X-axis represents business days of the month, the y-axis represents monthly spread.



[35] For Corn the period 1992 – 2004 is considered, for other commodities the period 1986 – 2004 is considered.

Table 6. Sources of Inventories data

Commodity	Source	Frequency and first date of observations
Crude oil	Energy Information administration, U.S. crude oil ending stocks non-SPR, thousand barrels.	Monthly, 1956 Weekly, 1990
Heating oil	Energy Information administration, U.S. total distillate stocks, thousand barrels.	Monthly, 1945 Weekly, 1982
Gasoline	Energy Information administration, U.S. total motor gasoline ending stocks, thousands of barrels.	Monthly, 1981 Weekly, 1983
Natural gas	Energy Information administration, U.S. total natural gas in underground storage (working gas), millions of cubic feet.	Monthly, 1975
Lean hogs	USDA, pork, Red meat and poultry beginning cold storage stocks, million pounds.	Monthly, 1917
Live cattle	USDA, beef, Red meat and poultry beginning cold storage stocks, million pounds.	Monthly, 1917
Cotton	ICE, Cotton N.2, Total Number of Bales in Designated Delivery Points as of the close of business every Wednesday from 1990-Aug 2002 Bloomberg, ICE cotton daily certified stockpiles, CTTOTOTL, bales.	Weekly, 1990
Cocoa	ICE, Historical cocoa warehouse stocks, thousand bags.	Monthly, 1986
Coffee	ICE, Historical coffee C warehouse stocks, bags.	Monthly, 1996
Copper	Bloomberg, COMEX copper stocks, COMXCOPR, tons.	Monthly, 1992
Corn	Bloomberg, USDA WASDE supply and use, total stocks of corn, millions of bushels, CUSEENDS.	Monthly, 1992
Wheat	Bloomberg, USDA WASDE Wheat total stocks, WUSETWES, millions of bushels.	Monthly, 1990
Soybeans	Bloomberg, USDA weekly stocks of grains in the selected elevators, soybeans, SOGRSOYB, thousands of bushels.	Weekly, 1974
Sugar	USDA, Sweetener market data, all total stocks, tons.	Monthly, 1991

The procedure of calculation of “Masters” implied positions (Masters & White, 2008):

There are two commodities Kansas Wheat (KW) and Feeder Cattle(FC) that are present in the GSCI and are absent in the DJ-UBSCI. By assuming that all Index trader positions reported in the Supplemental COT report of the CFTC are replicating the GSCI index for KW and FC, one can calculate the total dollars invested in the GSCI. For example, the formula for KW is the following:

Total dollars invested in the GSCI = (open interest of Kansas Wheat) X (price of KW) X (contract size of KW) / weight of KW in the GSCI

The price of KW is a price of the nearest-to-maturity futures contract, which is usually held by index traders. Then one can calculate the number of index trader positions for any commodity which is invested to replicate the GSCI. For instance, the following formula is for natural gas (NG):

Index trader positions of NG = (total dollars invested in the GSCI)x(weight of NG in the GSCI) / ((price of NG)(contract size of NG))*

The analogous procedure is applied for estimation of index trader positions which are replicated the DJ-UBSCI. In this case the reported position of Soybean Oil are used, which is included only in the DJ-UBSCI. To calculate the total index trader positions, the obtained positions associated with both indices are summed.

Figure 5. The spreading positions of wheat, 1996 – 2004, weekly, Datastream, CFTC.

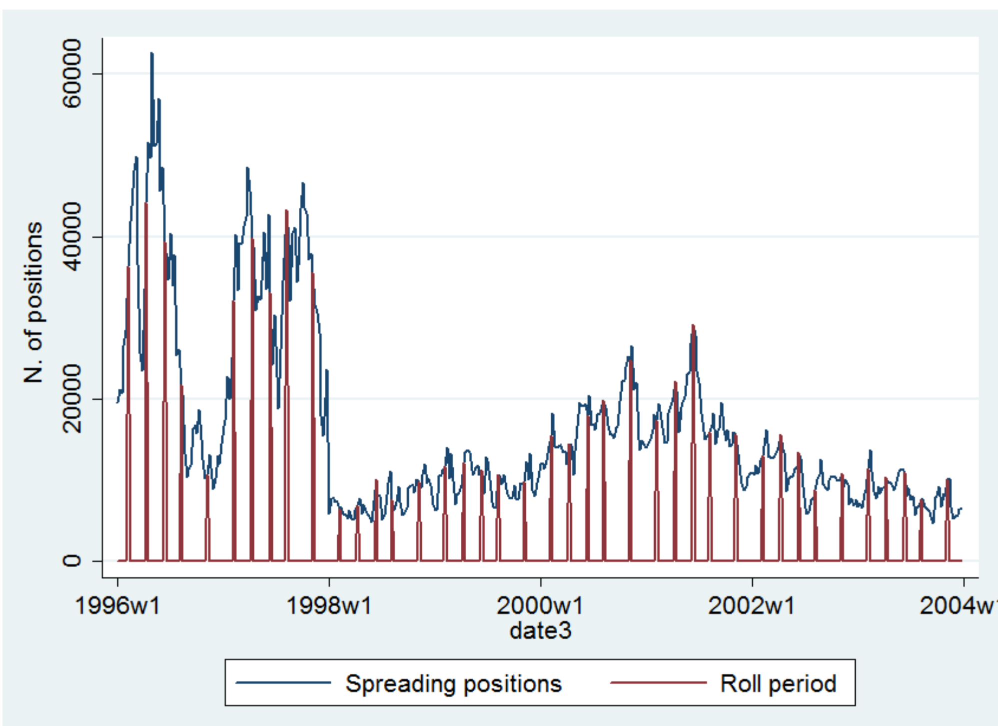


Table 7. The models used for forecasting IID positions.

The CFTC published the IID data about index trader positions since 2008 quarterly and since July 2010 monthly. The quarterly observations was joined linearly. The obtained series for long positions (IndexInvobs) and short positions (IndexInvobsS) were regressed against time, the “Masters” implied positions obtained from Kansas Wheat positions (long - IndexInv_KW, short - IndexInv_KWS), from Soybean oil (long - IndexInv_SO, short - IndexInv_SOS), from Feeder Cattle (long - IndexInv_FC, short - IndexInv_FCS), total open interest (TotOI), total reported short positions (TotS). Also changes of coefficients of “Masters” implied positions with a time are also included. For forecasting short positions, long positions predicted by the regression (IndexInvobs) are joined with reported long positions, the result is adjIndexInvobs. It is assumed that there is no constant in the model. The symbols *, **, and *** denote statistical significance at the 10%, 5% and 1% levels.

	Gasoline IndexInvobs	Gasoline IndexInvobsS	Heat. Oil IndexInvobs	Heat. Oil IndexInvobsS	Nat. Gas IndexInvobs	Nat. Gas IndexInvobsS	Crude Oil IndexInvobs	Crude Oil IndexInvobsS	Copper IndexInvobs	Copper IndexInvobsS
time	793.880***	198.629***	157.516***				1470.866***			
	(18.01)	(5.99)	(29.11)				(20.40)			
IndexInv_KW	0.368***		0.442***		2.548***		0.540***			
	(8.68)		(16.93)		(7.97)		(11.34)			
IndexInv_SO	0.801***									
	(8.45)									
IndexInv_KWt	-0.006***		-0.006***		-0.043***		-0.012***			
	(-8.63)		(-15.97)		(-6.56)		(-11.56)			
IndexInv_KWS		0.544*						-0.383*		
		(2.50)						(-2.12)		
IndexInv_KWSt		-0.012**		-0.004**		-0.012*		0.009*		0.021***
		(-3.03)		(-3.01)		(-2.20)		(2.57)		(5.84)
adjIndexInvobs		0.049*				0.225***		0.382***		
		(2.19)				(16.27)		(14.93)		
IndexInv_SOt			0.013***		0.011***		0.024***		0.023***	
			(12.07)		(3.97)		(7.02)		(18.25)	
IndexInv_SOS				-0.902*		-1.274**				1.011***
				(-2.36)		(-3.33)				(6.41)
IndexInv_SOSt				0.026**		0.037***		0.014*		
				(3.27)		(6.07)		(2.64)		
TotS				0.073***				-0.067***		
				(24.92)				(-6.19)		
IndexInv_FC					-3.059***		-0.497***		1.018***	
					(-6.52)		(-6.03)		(18.77)	
IndexInv_FCt					0.076***		0.008***		-0.018***	
					(8.25)		(5.00)		(-7.12)	
TotOI					0.191***					
					(4.75)					
IndexInv_FCS								-0.632***		
								(-5.48)		
IndexInv_FCSt								0.020***		0.036***
								(6.35)		(4.81)
N	64	64	64	64	62	62	64	64	64	64
adj. R^2	0.998	0.981	0.999	0.985	0.995	0.981	0.997	0.993	0.992	0.925

Table 8. The summary statistics of the spread of commodities over a short period.

	Gasoline	Heating oil	Natural gas	Crude oil	Copper	Cocoa	Coffee	Corn	Cotton	Soybeans	Sugar n. 11	Chicago Wheat	Lean Hogs	Live Cattle
mean[36]	0.0067	-0.0055	-0.0319	-0.0093	-0.0009	-0.0012	-0.0073	-0.0082	-0.0051	-0.0010	-0.0011	-0.0130	-0.0188	-0.0086
sd	0.0348	0.0130	0.0460	0.0174	0.0033	0.0062	0.0042	0.0108	0.0148	0.0106	0.0184	0.0184	0.0375	0.0140
min	-0.0692	-0.0442	-0.2949	-0.1014	-0.0095	-0.0213	-0.0194	-0.0233	-0.0305	-0.0133	-0.0517	-0.0319	-0.1049	-0.0386
max	0.1350	0.0481	0.0072	0.0146	0.0149	0.0193	0.0010	0.0338	0.0617	0.0580	0.0494	-0.0024	0.0731	0.0335
N. of obs.[37]	75	75	73	75	75	76	88	88	88	88	88	88	86	88

Table 9. The summary statistics of the spread of commodities over a long period.

	Gasoline	Heating oil	Natural gas	Crude oil	Copper	Cocoa	Coffee	Corn	Cotton	Soybeans	Sugar n. 11	Chicago Wheat	Lean Hogs	Live Cattle
mean	0.0031	0.0053	-0.0125	-0.0028	0.0017	-0.0057	-0.0068	-0.0085	-0.0031	-0.0014	-0.0037	-0.0056	-0.0070	0.0000
sd	0.0361	0.0396	0.0371	0.0095	0.0091	0.0084	0.0140	0.0147	0.0170	0.0093	0.0322	0.0173	0.0427	0.0171
min	-0.1416	-0.0566	-0.3078	-0.1014	-0.0105	-0.0287	-0.0486	-0.0244	-0.0869	-0.0133	-0.4055	-0.0319	-0.1549	-0.0386
max	0.1350	0.3029	0.0072	0.0146	0.0459	0.0385	0.0742	0.1304	0.1514	0.0580	0.1048	0.0912	0.1082	0.0486
N. of obs.	90	363	267	361	297	328	280	256	364	256	364	364	362	364

[36] sd – standard deviation, min – minimum, max - maximum

[37] N. of obs. – Number of observations. For Table 3 the number is corresponded with a period which was used to examine an impact of index trader positions on the spread

Table 10. Share of net index trader positions to the open interest.

	Gasoline	Heating oil	Natural gas	Crude oil	Copper	Cocoa	Coffee	Corn	Cotton	Soybeans	Sugar n. 11	Chicago Wheat	Lean Hogs	Live Cattle
mean	0.2665	0.2327	0.3005	0.3149	0.5026	0.1611	0.3316	0.3144	0.4146	0.3035	0.3138	0.4601	0.4314	0.4150
sd	0.0471	0.0297	0.1014	0.0528	0.0580	0.0446	0.0592	0.0424	0.0769	0.0564	0.0596	0.0531	0.0511	0.0560
min	0.1828	0.1678	0.1159	0.1937	0.3989	0.0577	0.2281	0.2304	0.2074	0.1567	0.1647	0.3155	0.3052	0.3018
max	0.3830	0.3109	0.4571	0.4012	0.6226	0.2602	0.5080	0.4145	0.6189	0.3906	0.4408	0.5646	0.5425	0.5304
N. of obs.	71	71	69	71	71	88	88	88	88	88	88	88	86	88
Tot OI[38], B$	41.26	41.96	36.92	153.42	14.95	4.71	9.79	46.28	7.75	53.96	18.62	18.73	9.11	17.44
DF[39]	0.04	0.03	0.59	0.86	0.59	0.05	0.02	0.11	0.01	0.28	0.08	0.01	0.04	0.16

[38] "Tot OI" is an average value of all open futures contracts calculated over 2012

[39] MacKinnon p-value for the Dickey-Fuller test

Table 11. Share of a sum of swap dealer and non-commercial spreading positions to the open interest.

	Gasoline	Heating Oil	Natural gas	Crude oil	Copper	Cocoa	Coffee	Corn	Cotton	Soybeans	Sugar	Wheat	Lean Hogs	Live Cattle
mean	0.1454	0.1913	0.4615	0.4336	0.1456	0.1271	0.1468	0.1541	0.1028	0.1651	0.1413	0.1985	0.2092	0.1646
sd	0.0266	0.0364	0.0415	0.0366	0.0369	0.0370	0.0344	0.0293	0.0286	0.0257	0.0355	0.0273	0.0309	0.0275
min	0.0643	0.1267	0.3683	0.3645	0.0659	0.0493	0.0608	0.0949	0.0404	0.1120	0.0608	0.1366	0.1260	0.0999
max	0.2043	0.3088	0.5697	0.5034	0.2358	0.2072	0.2115	0.2298	0.1863	0.2126	0.2347	0.2833	0.2843	0.2527
N. of obs.	76	78	78	78	84	83	84	84	84	84	84	84	85	84
SD spread share[40]	0.0544	0.0679	0.1225	0.1769	0.0411	0.0275	0.0198	0.0217	0.0220	0.0243	0.0463	0.0378	0.0160	0.0102
Correlation with Net CIT share[41]	-0.4315	0.1889	-0.7080	-0.5134	-0.0478	0.3725	-0.4061	-0.2392	-0.3738	-0.4707	-0.6062	-0.2528	-0.3274	-0.1256
DF[39]	0.0009	0.0399	0.0120	0.1223	0.0370	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Table 12. The correlation of spreading positions with Index Trader Long positions.

	Gasoline	Heating Oil	Natural gas	Crude oil	Copper	Cocoa	Coffee	Corn	Cotton	Soybeans	Sugar	Wheat	Lean Hogs	Live Cattle
NC spread	0.0132	-0.3664	-0.4749	-0.2251	0.7297	-0.1809	0.1625	0.5323	-0.2534	0.6155	0.2962	0.6498	0.2931	0.5416
SD spread	0.4076	0.0844	0.6590	-0.4778	0.6825	0.7295	0.2328	0.3583	0.6806	0.4595	0.5337	0.2430	0.6299	0.3679

NC spread – non-commercial spreading positions, SD spread – swap dealer spreading positions.

[40] The mean of swap dealer spreading position share to the open interest
[41] Correlation with Net index trader share to the open interest

Figure 6. The graph of “Tbill Adjusted” and “Yield Spread”.

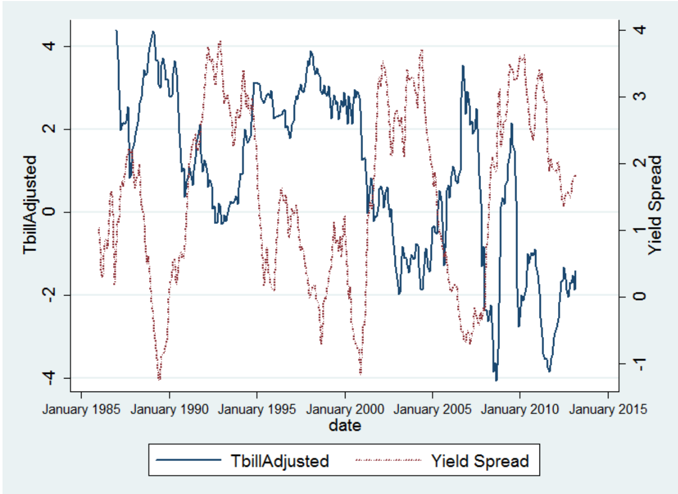


Figure 7. The graph of Cotton Stocks

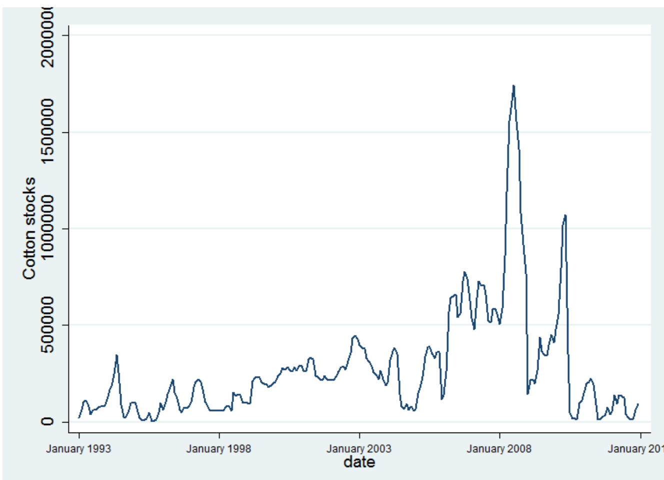


Table 13. Critical values for the F-test (Pesaran, Shin, & Smith, Bounds Testing Approaches to the Analysis of Level Relationships, 2001)

The critical values for the test of cointegration in a case without intercept in the ARDL model are computed by using the Monte-Carlo simulations (Pesaran et al (2001)). ”k” denotes the number of variables for the cointegration test. “I(0)” or “I(1)” designates two extreme cases when all variables are I(0) processes or I(1) processes correspondingly.

	90%		**95%**		**99%**	
k	**I(0)**	**I(1)**	**I(0)**	**I(1)**	**I(0)**	**I(1)**
1	2.44	3.28	3.15	4.11	4.81	6.02
2	2.17	3.19	2.72	3.83	3.88	5.30
3	2.01	3.01	2.45	3.63	3.42	4.84
4	1.90	3.01	2.26	3.48	3.07	4.44
5	1.81	2.93	2.14	3.34	2.82	4.21
6	1.75	2.87	2.04	3.24	2.66	4.05

Augmented Dickey-Fuller tests for the real interest rate and the yield spread

First, I chose the optimal number of lags by using the Akaike's criterion and the Shwarz's criterion (varsoc function in Stata). The optimal number for "TbillAdjustedusted" is 3 for the period 2007-2013. Then I carried out the ADF test[42]. I obtained the following results for "TbillAdjustedusted" :

Augmented Dickey-Fuller test for unit root Number of obs = 75

	Test Statistic	Interpolated Dickey-Fuller 1% Critical Value	5% Critical Value	10% Critical Value
Z(t)	-2.522	-3.545	-2.910	-2.590

MacKinnon approximate p-value for Z(t) = 0.1103

As cab be seen from Table, the hypothesis of the unit root can't be rejected. Then I use the first difference of this variable. I reproduced the same tests for the first difference and obtained that the unit root hypothesis is rejected at the 1% level:

Augmented Dickey-Fuller test for unit root Number of obs = 75

	Test Statistic	Interpolated Dickey-Fuller 1% Critical Value	5% Critical Value	10% Critical Value
Z(t)	-4.371	-3.545	-2.910	-2.590

MacKinnon approximate p-value for Z(t) = 0.0003

The ADF test with 1 lag for the yield spread for the period 2007-2013:

Augmented Dickey-Fuller test for unit root Number of obs = 75

	Test Statistic	Interpolated Dickey-Fuller 1% Critical Value	5% Critical Value	10% Critical Value
Z(t)	-2.029	-3.545	-2.910	-2.590

MacKinnon approximate p-value for Z(t) = 0.2739

The ADF test without lags for the first differnce of the yield spread for the period 2007-2013:

Dickey-Fuller test for unit root Number of obs = 75

	Test Statistic	Interpolated Dickey-Fuller 1% Critical Value	5% Critical Value	10% Critical Value
Z(t)	-7.276	-3.545	-2.910	-2.590

MacKinnon approximate p-value for Z(t) = 0.0000

As it can be seen,the hypothesis that the yield spread and the adjusted interest rate are non-stationary processes can't be rejected. However, the hypothesis is rejected for the first difference of these variables at 1% level.

[42] I did the test with a drift and as well as with a trend and without a drift and a trend. For all tests the hypothesis about a unit root can't be rejected.

Table 14. The impact of net index trader positions on the spread of energy commodities and copper during the Goldman roll.

The ARDL model (4) is estimated for 4 energy commodities and copper. The monthly observations on the 8th business day are used. The first difference of the spread of commodity futures (dSpread) is an independent variable. The dependent variables are the first differences of net index trader positions share in the total open interest (dNetCITshare), the yield spread (dYieldSpread), the hedging pressure (dHedgePressure), and the inflation adjusted interest rate (dTbillAdjusted), the total open interest and stocks normalized by their means for the period 2007-2013 (Stocks, TotalOI). The other dependent variables are the lagged variables of listed above and the lagged first difference of the spread (L.dSpread). First number in each cell is the OLS estimation of the coefficient for the corresponded variable. A number in brackets is the t-statistics. "N" is the number of observations. The results of the F-test for insignificant control variables are included under the "F-test of suppressed variables". The bottom line is the F-test for the presence of the cointegration of the level variables (lagged variables) in the right column model. The statistics compares with critical values of Pesaran, Table 8. If the value is higher the corresponding critical value, the absence of cointegration is rejected. In this case I insert "Yes" with the corresponding level of confidence. The symbols *, **, and *** denote statistical significance at the 10%, 5% and 1% levels.

	Heat. Oil	Heat. Oil	Nat. Gas	Nat. Gas	Crude Oil	Crude Oil	Copper	Copper
dNetCITshare	-0.006	0.010	-0.464**	-0.515**	0.041	0.017	-0.002	-0.002
	(-0.16)	(0.28)	(-2.35)	(-2.49)	(0.59)	(0.40)	(-0.61)	(-0.71)
dYieldSpread	-0.004		0.011		0.006		-0.001	
	(-1.64)		(1.24)		(0.80)		(-1.26)	
dStocks	0.008		-0.066**	-0.062**	-0.103**	-0.089**	-0.003**	-0.003***
	(0.43)		(-2.26)	(-2.61)	(-2.39)	(-2.57)	(-2.09)	(-3.95)
dHedgePressure	0.009		-0.000		0.018		0.007***	0.005***
	(0.50)		(-0.00)		(0.42)		(2.83)	(2.76)
dTbillAdjusted	-0.002		-0.010**	-0.010*	-0.005	-0.004	0.000	
	(-1.41)		(-2.18)	(-1.98)	(-1.66)	(-1.62)	(1.19)	
dTotalOI	0.002		-0.068		0.008		0.001	
	(0.14)		(-1.28)		(0.29)		(0.54)	
L.dSpread	-0.376***	-0.371***	0.444***	0.438***	0.147	0.087	-0.289**	-0.332**
	(-2.91)	(-3.21)	(5.03)	(6.99)	(0.68)	(0.46)	(-2.34)	(-2.43)
L.Spread	-0.437***	-0.402***	-0.819***	-0.801***	-0.487**	-0.377**	-0.637***	-0.552***
	(-2.75)	(-3.22)	(-6.93)	(-10.42)	(-2.48)	(-2.19)	(-3.89)	(-3.20)
L.NetCITshare	-0.067***	-0.070***	0.041	-0.024	-0.031	-0.011**	-0.001	-0.003**
	(-3.04)	(-3.66)	(0.80)	(-1.08)	(-0.98)	(-2.39)	(-0.79)	(-2.11)
L.HedgePressure	0.006		-0.000		0.011		0.003*	
	(0.27)		(-0.00)		(0.30)		(1.79)	
L.Stocks	0.018*	0.021***	-0.005		0.001		-0.002***	-0.001***
	(1.94)	(3.95)	(-0.42)		(0.04)		(-3.26)	(-2.79)
L.YieldSpread	-0.004***	-0.004***	-0.005		-0.001		-0.000	
	(-4.01)	(-4.03)	(-0.83)		(-0.65)		(-0.75)	
L.TotalOI	0.002		-0.033**	-0.027***	0.005		0.002**	0.002**
	(0.29)		(-2.60)	(-3.61)	(0.25)		(2.35)	(2.53)
L.TbillAdjusted	-0.001**	-0.001**	-0.008***	-0.007**	-0.001		-0.000	
	(-2.23)	(-2.05)	(-2.69)	(-2.83)	(-0.68)		(-0.33)	
N	70	70	68	68	70	70	70	70
adj. R^2	0.373	0.383	0.494	0.508	0.192	0.198	0.545	0.523
F test for suppressed variables	F(7, 56) = 0.86 Prob > F = 0.54		F(6, 53) = 0.70 Prob > F = 0.65		F(8, 56) = 1.05 Prob > F = 0.41		F(6, 56) = 1.52 Prob > F = 0.19	
F-test (cointegration)	F(5, 63) = 4.95 Yes***		F(4, 59) = 54.81 Yes***		F(2, 64) = 3.17 No		F(4, 62) = 2.96 No	

Table 15. The impact of net index trader positions on the spread of agricultural commodities during the Goldman roll.

The ARDL model (4) is estimated for 6 agricultural commodities. The dependent variable is the first difference of the spread. "NetCITshareroll" is a multiplication of "NetCITshare" and the roll period dummy (moving forward for the lagged variable). This variable represents the effect of index traders on the spread during the rolling period. The description of other variables is the same with Table 14. First number in each cell is the OLS estimation of the coefficient for the corresponded variable. A number in brackets is the t-statistics. "N" is the number of observations. The results of the F-test for insignificant control variables are included under the "F-test of suppressed variables". The bottom line is the F-test for the presence of the cointegration of the level variables (lagged variables) in the right column model. The statistics compares with critical values of Pesaran, Table 8. If the value is higher the corresponding critical value, the absence of cointegration is rejected. In this case I insert "Yes" with the corresponding level of confidence. The symbols *, **, and *** denote statistical significance at the 10%, 5% and 1% levels.

	Cocoa	Cocoa	Coffee	Coffee	Corn	Corn	Soybeans	Soybeans	Sugar	Sugar	Wheat	Wheat
dNetCITshareroll	-0.016	-0.030*	0.005	-0.002	0.068	0.021	-0.006	-0.000	-0.148**	-0.124**	-0.007	-0.018
	(-0.75)	(-1.74)	(0.63)	(-0.26)	(1.58)	(0.53)	(-0.19)	(-0.03)	(-2.09)	(-2.01)	(-0.28)	(-0.90)
dYieldSpread	0.001		-0.000		0.003		0.004*	0.004	-0.004		0.002	
	(0.77)		(-0.60)		(1.41)		(1.82)	(1.55)	(-0.95)		(0.67)	
dStocks	-0.001		-0.014***	-0.012***	-0.005		-0.003**	-0.001	0.006		0.003	
	(-0.20)		(-2.90)	(-3.56)	(-1.48)		(-2.05)	(-1.24)	(0.74)		(0.38)	
dHedgePressure	0.005		0.008***	0.008***	0.021	0.018*	0.011		0.054**	0.040*	0.018	
	(0.95)		(3.81)	(5.02)	(1.59)	(1.70)	(1.23)		(2.03)	(1.70)	(1.36)	
dTbillAdjusted	0.000		0.000		-0.000		0.002**	0.001*	-0.001		-0.000	
	(0.19)		(1.51)		(-0.26)		(2.05)	(1.91)	(-0.34)		(-0.61)	
dTotalOI	0.002		0.002		0.005		-0.009		-0.017		-0.001	
	(0.46)		(1.12)		(0.47)		(-1.06)		(-1.06)		(-0.10)	
L.dSpread	-0.286**	-0.387***	-0.313***	-0.317***	0.015	0.026	0.262	0.176	-0.228**	-0.185*	-0.184	-0.328**
	(-2.49)	(-4.02)	(-3.16)	(-3.92)	(0.18)	(0.32)	(1.34)	(0.83)	(-2.00)	(-1.78)	(-1.31)	(-2.30)
L.Spread	-0.386***	-0.169***	-0.107	-0.049***	-0.626***	-0.616***	-0.554***	-0.347**	-0.160*	-0.154**	-0.384***	-0.122***
	(-3.15)	(-2.66)	(-1.40)	(-3.19)	(-9.97)	(-10.25)	(-3.03)	(-2.22)	(-1.81)	(-2.25)	(-2.84)	(-2.71)
L.NetCITshareroll	0.000	-0.003	-0.002***	-0.002***	-0.004	-0.004	-0.002	-0.006*	-0.006	-0.012	-0.005*	-0.006***
	(0.15)	(-1.30)	(-2.98)	(-3.42)	(-0.79)	(-0.89)	(-0.64)	(-1.82)	(-0.63)	(-1.62)	(-1.94)	(-2.79)
L.HedgePressure	0.000		0.000		-0.003		0.012*		0.032	0.008	-0.000	
	(0.08)		(0.30)		(-0.21)		(1.78)		(1.53)	(1.00)	(-0.07)	
L.Stocks	-0.001		-0.001		-0.009***	-0.007***	-0.002**		-0.006		0.001	
	(-0.46)		(-1.09)		(-5.01)	(-6.86)	(-2.02)		(-1.32)		(0.28)	
L.TotalOI	-0.001		0.000		0.003		-0.002		0.003		-0.003	
	(-0.42)		(0.42)		(0.92)		(-0.95)		(0.50)		(-0.83)	
L.YieldSpread	0.000		-0.000		0.001**	0.001***	0.001**	0.000	-0.001		-0.001*	
	(0.58)		(-0.37)		(2.29)	(3.35)	(2.10)	(1.51)	(-0.44)		(-1.75)	
L.TbillAdjusted	-0.000		-0.000		-0.000		0.000		0.000		-0.000	
	(-0.92)		(-0.11)		(-0.79)		(0.74)		(0.63)		(-1.24)	
N	86	87	86	87	83	83	86	86	83	86	86	87
adj. R^2	0.260	0.289	0.424	0.430	0.578	0.566	0.189	0.163	0.166	0.173	0.241	0.223
F test for suppressed variables	F(10, 72) = 0.72 Prob > F = 0.70		F(8, 72) = 0.93 Prob > F = 0.50		F(7, 69) = 1.32 Prob > F = 0.26		F(6, 72) = 1.14 Prob > F = 0.35		F(8, 69) = 0.81 Prob > F = 0.60		F(10, 72) = 1.10 Prob > F = 0.38	
F-test (cointegration)	F(2, 83) = 3.59 Yes*		F(2, 81) = 7.12 Yes***		F(4, 76) = 27.19 Yes***		F(3, 78) = 3.51 Yes*		F(3, 80) = 2.32 No		F(2, 83) = 4.98 Yes**	

Table 16. The impact of net index trader positions on the spread of commodities with the strong seasonality during the Goldman roll.

The ARDL model (4) is estimated for four commodities with the strong seasonality. The dependent variable is the first difference of the spread. The monthly observations on the 8th business day are used. All variables are described in the same way as Table 15. First number in each cell is the OLS estimation of the coefficient for the corresponded variable. Monthly dummies are included where the number corresponds with the order of the month (m1 - January). The t-statistics of monthly dummy variables is suppressed. A number in brackets is the t-statistics. The symbols *, **, and *** denote statistical significance at the 10%, 5% and 1% levels.

	Gasoline	Gasoline	Lean Hogs	Lean Hogs	Live Cattle	Live Cattle	Cotton	Cotton
dNetCITshare	0.007	0.010						
	(0.12)	(0.18)						
dNetCITshareroll			-0.034	-0.038	-0.123**	-0.093**	-0.123	-0.066
			(-0.47)	(-0.59)	(-2.33)	(-2.01)	(-1.36)	(-0.96)
dYieldSpread	0.018**	0.016**	-0.005		-0.005		-0.000	
	(2.38)	(2.13)	(-0.79)		(-1.53)		(-0.00)	
dStocks	-0.057		0.027		0.018		-0.009***	-0.005*
	(-1.47)		(0.71)		(0.74)		(-2.89)	(-1.92)
dHedgePressure	0.002		0.019		0.000		-0.011	
	(0.07)		(0.85)		(0.02)		(-0.93)	
dTbillAdjusted	-0.004*		0.001		-0.001		0.002	
	(-1.74)		(0.24)		(-0.50)		(1.46)	
dTotalOI	0.002		0.040*	0.039*	-0.036**	-0.020	-0.005	
	(0.07)		(1.74)	(1.95)	(-2.01)	(-1.38)	(-0.36)	
L.dSpread	0.106	0.118	-0.049	-0.068	0.236*	0.227**	-0.089	-0.242*
	(0.90)	(1.13)	(-0.38)	(-0.57)	(1.84)	(2.07)	(-0.66)	(-1.81)
L.Spread	-0.899***	-0.859***	-0.529***	-0.399***	-0.561***	-0.529***	-0.546***	-0.244**
	(-4.75)	(-4.42)	(-3.75)	(-3.60)	(-4.47)	(-5.46)	(-3.81)	(-2.53)
L.NetCITshare	-0.136**	-0.099***						
	(-2.53)	(-2.72)						
L.NetCITshareroll			-0.054	-0.068*	-0.055**	-0.067***	-0.036	-0.030
			(-1.21)	(-1.85)	(-2.14)	(-3.17)	(-1.19)	(-1.05)
L.HedgePressure	-0.041		0.020		0.015		-0.007	
	(-0.85)		(1.04)		(0.85)		(-0.67)	
L.Stocks	0.062**	0.043***	-0.007		0.002		-0.004**	
	(2.42)	(4.57)	(-0.45)		(0.14)		(-2.00)	
L.TotalOI	-0.009		0.004		-0.011		0.016**	
	(-0.45)		(0.24)		(-1.07)		(2.22)	
L.TbillAdjusted	0.002		-0.000		0.000		0.000	
	(1.65)		(-0.15)		(0.49)		(0.37)	
L.YieldSpread[43]			-0.000		0.000		0.003***	
			(-0.27)		(0.58)		(2.73)	
m1	-0.045***	-0.045***	0.017	0.019	0.019	0.024**	0.012	0.004
m2	0.016	0.019**	-0.027***	-0.029***	0.024***	0.023***	-0.012**	-0.014**
m3	-0.005	-0.001	-0.005	0.003	0.030***	0.032***	-0.003	-0.006
m4	0.003	0.007	0.032***	0.033***	-0.003	-0.006	0.010	0.005
m5	0.009	0.013**	0.041	0.040**	0.029**	0.032***	-0.002	-0.005
m6	0.014***	0.015***	0.038	0.035	-0.015***	-0.016***	0.012	0.005
m7	0.064***	0.066***	0.083***	0.078***	0.022	0.027***	-0.007	-0.011**
m8	0.010	0.009	0.009	-0.005	0.009**	0.011***	0.012	0.004
m9	0.021	0.028***	0.039	0.036	0.019	0.026**	-0.009**	-0.014**
m10	0.009	0.013**	-0.027***	-0.035***	-0.001	-0.000	-0.009	-0.012**
m11	-0.000	0.004	0.006	0.010	0.022	0.027**	-0.012**	-0.013**
constant	-0.016	-0.028**	-0.009	-0.006	0.002	-0.006**	-0.013	0.007*
	(-0.96)	(-2.53)	(-0.41)	(-0.88)	(0.09)	(-2.05)	(-1.49)	(1.79)
N	70	70	85	85	86	87	86	87
adj. R^2	0.911	0.903	0.827	0.833	0.766	0.772	0.235	0.228
F test for suppressed variables	F(7, 45) = 1.54 Prob > F = 0.18		F(9, 59) = 0.71 Prob > F = 0.70		F(9, 60) = 0.83 Prob > F = 0.59		F(9, 60) = 1.22 Prob > F = 0.30	
F-test (cointegration)	F(3, 52) = 9.32 Yes***		F(2, 68) = 7.23 Yes***		F(2, 70) = 16.21 Yes***		F(2, 70) = 4.40 Yes***	

[43] Yield spread is excluded from the Gasoline equation. It has a high correlation with net index trader share.

Table 17. The long-run effect of the net index trader position share on the spread.

The model (4) is estimated by the OLS. The dependent variable is the spread. The OLS regressions are estimated only for those commodities that have statistically significant coefficient of the ”NetCITshare” and reject the hypothesis about the absence of the cointegration. The estimation is carried out for the variables that are significant in the right column of each commodity in Tables 14, 15 and 16. The order of the lagged variables was chosen by the SC and AIC. The long-run marginal effect is calculated by the formula (5). The impact of one standard deviation of the net index trader share (“1 sd impact”) is computed as a multiplication of the marginal effect and the standard deviation of “NetCITshare” from Table 10 of Appendix. The standard deviation of the spread of the corresponding commodity is attached in the bottom line. Monthly dummies are included where the number corresponds with the order of the month (m1 - January). Adjustment coefficient is the estimated coefficient δ in the model (6). It represents the speed of convergence to the equilibrium. The symbols *, **, and *** denote statistical significance at the 10%, 5% and 1% levels.

	Gasoline	Heating Oil	Lean Hogs	Live Cattle	Coffee	Soybeans	Wheat
L.Spread	0.308**	0.248*	0.537***	0.716***	0.670***	0.711***	0.545***
	(2.58)	(2.24)	(5.61)	(6.73)	(8.49)	(5.78)	(3.57)
L2.Spread		0.416***		-0.240**	0.286***		0.331**
		(3.75)		(-2.27)	(3.40)		(2.27)
NetCITshare	-0.069**	-0.066***					
	(-2.25)	(-3.84)					
NetCITshareroll			-0.076**	-0.066***	-0.002***	-0.006***	-0.006***
			(-2.20)	(-3.46)	(-2.83)	(-2.93)	(-3.06)
Stocks	-0.086*	0.020***					
	(-1.73)	(4.06)					
YieldSpread		-0.004***				0.001	
		(-4.55)				(1.41)	
TbillAdjusted		-0.001*					
		(-2.48)					
HedgePressure							
m1	-0.042***		0.020	0.024***			
m2	0.027**		-0.028***	0.023***			
m3	0.008		0.005	0.032***			
m4	0.003		0.030***	-0.006			
m5	0.009		0.041**	0.032***			
m6	0.010		0.038**	-0.014***			
m7	0.062***		0.082***	0.028***			
m8	0.000		-0.004	0.011**			
m9	0.009		0.043**	0.026***			
m10	0.006		-0.034***	0.001			
m11	0.001		0.014	0.027**			
constant	-0.015		-0.006	-0.006**			
	(-1.51)		(-1.01)	(-2.29)			
N	70	71	86	87	87	87	87
adj. R^2	0.861	0.766	0.860	0.760	0.980	0.589	0.885
Long-run marginal effect	-0.109	-0.197	-0.164	-0.127	-0.046	-0.022	-0.045
1 sd effect	-0.473%	-0.583%	-0.838%	-0.710%	-0.269%	-0.129%	-0.255%
Sd of spread	3.483%	1.296%	3.750%	1.401%	0.425%	1.059%	0.615%
Adjustment coefficient	1.188***	1.150***	0.752***	0.765***	-0.146	0.870*	0.506*

Table 18. The impact of spreading positions on the spread of energy commodities and copper before the Goldman roll on the 1st business day of the month.

The ARDL model (4) is estimated for 4 energy commodities and copper. The dependent variable is the first difference of the spread. The monthly observations on the 1th business day are used. "Spreadingshare" is a share of spreading positions of non-commercials and swap dealers to the open interest. "Spreadingshareroll" is a multiplication of "Spreadingshare" and the roll period dummy (moving forward for the lagged variable). This variable represents the effect of spreading positions on the spread during the rolling period. "dSpreadingshareroll" is the first difference of the variable. "L.Spreadingshareroll" is the first lag of the variable. The description of other variables coincides with the previous models. First number in each cell is the OLS estimation of the coefficient for the corresponded variable. A number in brackets is the t-statistics. "N" is the number of observations. The results of the F-test for insignificant control variables are included under the "F-test of suppressed variables". The bottom line is the F-test for the presence of the cointegration of the level variables (lagged variables) in the right column model. The statistics compares with critical values of Pesaran, Table 8. If the value is higher the corresponding critical value, the absence of cointegration is rejected. In this case I insert "Yes" with the corresponding level of confidence. The symbols *, **, and *** denote statistical significance at the 10%, 5% and 1% levels.

	Heat.Oil	Heat. Oil	Nat. Gas	Nat. Gas	Crude Oil	Crude Oil	Copper	Copper
dSpreadingshare	-0.073^{**}	-0.066^{***}	0.054	0.063	0.085^{**}	0.086^{***}	-0.013^{*}	-0.014^{**}
	(-2.94)	(-2.86)	(0.57)	(0.83)	(2.92)	(3.50)	(-1.97)	(-2.35)
dYiledSpread	0.001		0.002		0.005^{*}	0.006^{***}	-0.000	
	(0.68)		(0.27)		(1.84)	(2.38)	(-0.40)	
dStocks	-0.009		-0.005		-0.032		-0.002	-0.002^{*}
	(-0.71)		(-0.31)		(-1.22)		(-1.17)	(-1.78)
dHedgePressure	0.005		0.017		0.057^{**}	0.062^{***}	0.007^{***}	0.007^{***}
	(0.36)		(0.26)		(2.85)	(3.51)	(3.05)	(3.89)
dTbillAdjusted	-0.003^{**}	-0.003^{***}	-0.000		-0.002	-0.002	-0.000	
	(-3.21)	(-3.79)	(-0.05)		(-1.31)	(-1.53)	(-0.11)	
dTotalOI	-0.009		-0.026		-0.048^{**}	-0.056^{***}	0.000	
	(-1.02)		(-0.86)		(-2.55)	(-3.70)	(0.20)	
L.dSpread	0.020	-0.028	0.439^{***}	0.394^{***}	0.110	0.062	-0.149	-0.146
	(0.19)	(-0.30)	(4.45)	(5.63)	(0.81)	(0.62)	(-1.10)	(-1.10)
L.Spread	-0.805^{***}	-0.679^{***}	-0.474^{**}	-0.415^{***}	-0.373^{**}	-0.275^{***}	-0.536^{***}	-0.520^{***}
	(-5.84)	(-5.78)	(-3.38)	(-5.14)	(-2.72)	(-3.51)	(-3.74)	(-3.79)
L.Spreadingshare	-0.054^{***}	-0.052^{***}	0.006	-0.015^{*}	-0.030	-0.040^{***}	0.010^{*}	0.006^{*}
	(-4.43)	(-5.10)	(0.25)	(-1.80)	(-1.54)	(-3.65)	(1.68)	(1.81)
L.HedgePressure	0.017		-0.055		0.007		0.004^{**}	0.003^{**}
	(1.16)		(-0.96)		(0.46)		(2.28)	(2.18)
L.Stocks	-0.005		-0.007		-0.004		-0.001	-0.001
	(-0.97)		(-0.63)		(-0.31)		(-1.47)	(-1.49)
L.YieldSpread	-0.004^{***}	-0.004^{***}	-0.007^{*}	-0.004^{**}	-0.001		-0.000^{**}	-0.000^{**}
	(-5.12)	(-5.26)	(-1.78)	(-2.20)	(-0.94)		(-2.20)	(-2.06)
L.TotalOI	0.014^{***}	0.011^{***}	-0.008		0.015	0.015^{***}	-0.000	
	(3.35)	(4.66)	(-0.52)		(1.34)	(3.19)	(-0.20)	
L.TbillAdjusted	-0.002^{***}	-0.002^{***}	-0.004^{**}	-0.003^{**}	-0.000		0.000	
	(-4.26)	(-4.29)	(-2.52)	(-2.15)	(-0.38)		(1.12)	
N	74	74	68	69	74	74	75	75
adj. R^2	0.398	0.410	0.328	0.363	0.429	0.445	0.434	0.454
F test for suppressed variables	$F(6, 60) = 0.77$ Prob > F = 0.60		$F(8, 54) = 0.72$ Prob > F = 0.67		$F(6, 60) = 0.94$ Prob > F = 0.48		$F(5, 61) = 0.52$ Prob > F = 0.77	
F-test (cointegration)	$F(5, 66) = 7.59$ Yes***		$F(4, 63) = 7.56$ Yes***		$F(3, 65) = 8.63$ Yes***		$F(5, 66) = 3.03$ Yes*	

Table 19. The impact of spreading positions on the spread of agricultural commodities before the Goldman roll on the 1st business day of the month.

The ARDL model (4) is estimated for 7 agricultural commodities. The dependent variable is the first difference of the spread. The monthly observations on the 1th business day are used. "Spreadingshareroll" is a share of spreading positions of non-commercials and swap dealers to the open interest. "dSpreadingshare" is the first difference of the variable. "L.Spreadingshare" is the first lag of the variable. The description of other variables coincides with the previous models. First number in each cell is the OLS estimation of the coefficient for the corresponded variable. A number in brackets is the t-statistics. "N" is the number of observations. The results of the F-test for insignificant control variables are included under the "F-test of suppressed variables". The bottom line is the F-test for the presence of the cointegration of the level variables (lagged variables) in the right column model. The statistics compares with critical values of Pesaran, Table 8. If the value is higher the corresponding critical value, the absence of cointegration is rejected. In this case I insert "Yes" with the corresponding level of confidence. The symbols *, **, and *** denote statistical significance at the 10%, 5% and 1% levels.

	Cocoa	Cocoa	Coffee	Coffee	Corn	Corn	Soybeans	Soybeans	Sugar	Sugar	Wheat	Wheat
dSpreadingshareroll	-0.006	-0.006	-0.010	-0.011	-0.095**	-0.091**	-0.019	0.009	-0.009	-0.025	0.026	0.031
	(-0.35)	(-0.38)	(-0.89)	(-1.09)	(-2.14)	(-2.57)	(-0.46)	(0.30)	(-0.08)	(-0.23)	(0.95)	(1.25)
dYieldSpread	0.001		-0.000		0.001		0.004		-0.003		0.005	
	(1.14)		(-0.67)		(0.65)		(1.48)		(-0.65)		(0.81)	
dStocks	-0.005	-0.005	-0.017**	-0.017**	-0.003		-0.002		0.006		0.008	
	(-1.40)	(-1.24)	(-2.09)	(-2.20)	(-0.91)		(-1.19)		(0.67)		(0.79)	
dHedgePressure	0.012*	0.009	0.008**	0.007**	-0.012		0.007		0.087***	0.067***	0.028*	0.024**
	(1.70)	(1.55)	(2.04)	(2.37)	(-0.93)		(0.66)		(2.81)	(2.64)	(1.74)	(2.31)
dTbillAdjusted	-0.000		0.000		-0.001		0.002**		0.000		-0.001	
	(-0.05)		(0.17)		(-1.18)		(2.05)		(0.02)		(-0.81)	
dTotalOI	-0.005		0.001		0.031***	0.022***	-0.003		-0.015		-0.001	
	(-1.13)		(0.33)		(2.69)	(2.87)	(-0.37)		(-0.75)		(-0.12)	
L.dSpread	-0.211**	-0.253***	-0.226**	-0.212**	0.029	0.076	0.239	0.180	-0.162	-0.111	-0.129	-0.236
	(-2.15)	(-2.99)	(-2.06)	(-2.07)	(0.38)	(1.06)	(1.43)	(0.97)	(-1.32)	(-1.02)	(-0.78)	(-1.49)
L.Spread	-0.396***	-0.349***	-0.258	-0.260***	-0.524***	-0.531***	-0.678***	-0.440**	-0.201**	-0.176**	-0.618***	-0.453***
	(-2.85)	(-3.99)	(-1.57)	(-3.09)	(-7.78)	(-8.52)	(-3.64)	(-2.56)	(-2.06)	(-2.43)	(-3.41)	(-3.87)
L.Spreadingshareroll	-0.013**	-0.016***	-0.007**	-0.006***	-0.013	-0.009	-0.017**	-0.022***	-0.018	-0.033*	-0.025***	-0.026***
	(-2.28)	(-3.93)	(-2.15)	(-2.91)	(-1.76)	(-1.30)	(-2.56)	(-3.49)	(-0.79)	(-1.81)	(-2.93)	(-4.50)
L.HedgePressure	0.000		0.001		-0.012		0.020**		0.037	0.014	0.001	
	(0.14)		(0.44)		(-1.03)		(2.40)		(1.34)	(1.61)	(0.12)	
L.Stocks	0.001		-0.002**	-0.002**	-0.008***	-0.004***	-0.002		-0.004		0.001	
	(0.31)		(-2.01)	(-2.21)	(-4.12)	(-5.33)	(-1.77)		(-0.88)		(0.13)	
L.TotalOI	-0.001		0.000		0.005		-0.003*		0.001		-0.003	
	(-0.68)		(0.35)		(1.60)		(-1.73)		(0.14)		(-0.46)	
L.YieldSpread	0.000		-0.000		0.000		0.002**	0.001**	-0.000		-0.002**	-0.002***
	(0.41)		(-0.41)		(0.80)		(2.00)	(2.01)	(-0.20)		(-2.06)	(-3.01)
L.TbillAdjusted	-0.000		-0.000		-0.000	-0.000**	0.001		0.000		-0.001	
	(-0.52)		(-0.05)		(-0.92)	(-2.05)	(0.87)		(0.51)		(-1.19)	
N	81	82	81	82	77	77	81	82	78	81	81	82
adj. R^2	0.372	0.416	0.323	0.382	0.545	0.550	0.264	0.174	0.129	0.174	0.421	0.426
F test for suppressed variables	F(8, 67) = 0.80 Prob > F = 0.61		F(8, 66) = 1.16 Prob > F = 0.34		F(7, 63) = 0.88 Prob > F = 0.53		F(9, 67) = 1.33 Prob > F = 0.24		F(7, 64) = 0.20 Prob > F = 0.98		F(8, 67) = 0.77 Prob > F = 0.63	
F-test (cointegration)	F(2, 76) = 11.66 Yes***		F(3, 75) = 10.61 Yes***		F(4, 70) = 19.43 Yes***		F(3, 77) = 4.58 Yes**		F(3, 75) = 3.35 Yes*		F(3, 76) = 7.14 Yes***	

Table 20. The impact of spreading positions on the spread of commodities with the strong seasonality before the Goldman roll.

The ARDL model (4) is estimated for four commodities with the strong seasonality. The dependent variable is the first difference of the spread. The monthly observations on the 1st business day are used. All variables are described in the same way as Table 15. First number in each cell is the OLS estimation of the coefficient for the corresponded variable. Monthly dummies are included where the number corresponds with the order of the month (m1 - January). The t-statistics of monthly dummy variables is suppressed. A number in brackets is the t-statistics. "N" is the number of observations. The results of the F-test for insignificant control variables are included under the "F-test of suppressed variables". The bottom line is the F-test for the presence of the cointegration of the level variables (lagged variables) in the right column model. The statistics compares with critical values of Pesaran, Table 8. If the value is higher the corresponding critical value, the absence of cointegration is rejected. In this case I insert "Yes" with the corresponding level of confidence. The symbols *, **, and *** denote statistical significance at the 10%, 5% and 1% levels.

	Lean Hogs	Lean Hogs	Live Cattle	Live Cattle	Cotton	Cotton	Gasoline	Gasoline
dSpreadingshareroll	-0.161	-0.077	0.037	0.018	-0.113	-0.118*	0.067	0.028
	(-1.09)	(-0.60)	(0.50)	(0.29)	(-1.68)	(-1.94)	(0.92)	(0.46)
dYieldSpread	-0.003		-0.004		-0.003		0.002	
	(-0.44)		(-1.14)		(-1.15)		(0.29)	
dStocks	0.013		0.024		-0.003		-0.078*	-0.077*
	(0.30)		(0.88)		(-1.55)		(-1.80)	(-1.98)
dHedgePressure	0.002		-0.014		0.007		0.040	
	(0.09)		(-0.64)		(0.88)		(1.40)	
dTbillAdjusted	0.004		0.001		0.002		-0.000	
	(1.46)		(0.37)		(1.02)		(-0.21)	
dTotalOI	0.077***	0.059**	0.009		0.008		-0.004	
	(3.01)	(2.87)	(0.48)		(0.77)		(-0.30)	
L.dSpread	-0.135	-0.240*	0.050	0.019	-0.060	-0.052	0.010	0.061
	(-1.03)	(-2.15)	(0.32)	(0.15)	(-0.52)	(-0.51)	(0.07)	(0.46)
L.Spread	-0.678***	-0.461***	-0.494**	-0.425***	-0.286***	-0.193**	-0.879***	-0.924***
	(-4.02)	(-3.94)	(-3.03)	(-3.68)	(-2.93)	(-2.10)	(-5.19)	(-5.69)
L.Spreadingshareroll	-0.143	-0.060	0.036	0.013	0.035	0.024	-0.017	-0.050
	(-1.39)	(-0.78)	(0.57)	(0.26)	(0.69)	(0.41)	(-0.17)	(-0.78)
L.HedgePressure	0.037*	0.035*	0.013		-0.001		0.070**	0.039*
	(1.78)	(2.35)	(0.55)		(-0.13)		(2.18)	(1.86)
L.Stocks	0.008		0.011		-0.002	-0.000	0.044***	0.037***
	(0.47)		(0.56)		(-1.24)	(-0.24)	(3.19)	(4.25)
L.TotalOI	0.031		-0.006		0.013*		-0.005***	-0.006***
	(1.50)		(-0.41)		(1.82)		(-3.80)	(-4.76)
L.YieldSpread	0.001		0.001		0.002*	0.001*	-0.016	
	(0.78)		(0.69)		(2.25)	(1.83)	(-1.07)	
L.TbillAdjusted	0.003		0.000		0.000		-0.000	
	(1.51)		(0.32)		(0.65)		(-0.08)	
m1	0.028	0.010	-0.011	-0.007	-0.007	-0.007**	-0.004	-0.004
m2	-0.022***	-0.024***	0.021***	0.022***	-0.012	-0.010	0.010	0.008
m3	0.003	-0.014	-0.001	0.002	-0.006	-0.007	0.009	0.007
m4	0.030***	0.034***	-0.002	-0.004	-0.012	-0.011	0.007	0.005
m5	0.040	0.020	-0.004	-0.002	-0.005	-0.004	0.011	0.010
m6	0.048*	0.025	-0.013***	-0.015***	-0.009	-0.008	0.023***	0.020***
m7	0.096***	0.073***	-0.012	-0.009	-0.007	-0.007	0.066***	0.065***
m8	0.027	0.014	0.010	0.011**	-0.009	-0.009**	0.013	0.013
m9	0.052*	0.019	-0.007	-0.003	-0.010*	-0.010**	0.017	0.018
m10	-0.021	-0.033***	0.003	0.003	-0.014*	-0.013**	0.012*	0.011**
m11	0.014	-0.009	-0.009	-0.007	-0.014	-0.012	0.000	0.002
constant	-0.058**	-0.011	-0.012	-0.004	-0.007	0.006*	-0.037*	-0.033***
	(-2.13)	(-1.53)	(-0.61)	(-1.36)	(-0.85)	(2.05)	(-2.32)	(-2.94)
N	80	80	81	82	81	82	74	74
adj. R^2	0.834	0.844	0.660	0.674	0.334	0.343	0.863	0.868
F test for suppressed variables	F(9, 59) = 0.42 Prob > F = 0.92		F(10, 55) = 0.71 Prob > F = 0.71		F(8, 55) = 0.96 Prob > F = 0.48		F(6, 48) = 1.05 Prob > F = 0.41	
F-test (cointegration)	F(2, 68) = 5.21 Yes**		F(2, 66) = 6.80 Yes***		F(4, 64) = 1.37 No		F(5, 54) = 5.63 Yes***	

Table 21. The long-run effect of the spreading position share on the spread.

The OLS regressions are estimated only for those commodities that have statistically significant coefficient of ”Spreadingshare” and reject the hypothesis about the absence of the cointegration. The estimation is carried out for the variables that are significant in the right column of each commodity in Tables 18, 19 and 20. The order of the lagged variables was chosen by the SC and AIC. The long-run marginal effect is calculated by the formula (5). The impact of one standard deviation of the net index trader share (“1 sd impact”) is computed as a multiplication of the marginal effect and the standard deviation of ”Spreadingshare” from Table 11 of Appendix. The standard deviation of the spread of the corresponding commodity is attached in the bottom line. Adjustment coefficient is the estimated coefficient δ in the model (6). It represents the speed of convergence to the equilibrium. The symbols *, **, and *** denote statistical significance at the 10%, 5% and 1% levels.

	Heat.Oil	Nat. Gas	Crude Oil	Cocoa	Coffee	Soybeans	Sugar	Wheat
L.Spread	0.319***	1.001***	0.819***	0.405***	0.712***	0.619***	0.779***	0.493**
	(3.30)	(7.79)	(12.87)	(4.45)	(8.37)	(5.11)	(8.73)	(3.04)
L2.Spread		-0.399***		0.248**				
		(-7.02)		(2.62)				
Spreadingshare	-0.045***	-0.017**	-0.020*					
	(-4.63)	(-2.66)	(-1.68)					
TotalOI	0.010***		0.007					
	(4.44)		(1.48)					
Spread	-0.004***	-0.003*				0.001		-0.002**
	(-5.99)	(-1.98)				(1.56)		(-2.55)
TbillAdjustedusted	-0.002***	-0.002*						
	(-4.45)	(-1.92)						
Spreadingshareroll				-0.016***	-0.006***	-0.018***	-0.039**	-0.025***
				(-4.36)	(-3.78)	(-3.53)	(-2.59)	(-3.55)
Stocks					-0.002**			
					(-2.47)			
HedgePressure							0.018**	
							(2.06)	
N	75	68	77	83	83	82	82	82
adj. R^2	0.815	0.823	0.781	0.627	0.965	0.488	0.645	0.764
Long-run marginal effect	-0.066	-0.043	-0.110	-0.046	-0.021	-0.047	-0.176	-0.049
1 sd effect	-0.240%	-0.177%	-0.405%	-0.170%	-0.072%	-0.121%	-0.627%	-0.135%
Sd of spread	1.296%	4.599%	1.739%	0.620%	0.425%	1.059%	1.841%	0.615%
Adjustment coefficient	0.995***	1.021***	1.695***	0.126	0.026	0.029	0.314	0.522*

Table 22. The impact of spreading positions on the spread of energy commodities and copper during the Goldman roll on the 8th business day of the month.

The ARDL model (4) is estimated for 4 energy commodities and copper. The dependent variable is the first difference of the spread. The monthly observations on the 8th business day are used. The description of other variables coincides with the previous models. First number in each cell is the OLS estimation of the coefficient for the corresponded variable. A number in brackets is the t-statistics. "N" is the number of observations. The results of the F-test for insignificant control variables are included under the "F-test of suppressed variables". The bottom line is the F-test for the presence of the cointegration of the level variables (lagged variables) in the right column model. The statistics compares with critical values of Pesaran, Table 8. If the value is higher the corresponding critical value, the absence of cointegration is rejected. In this case I insert "Yes" with the corresponding level of confidence. The symbols *, **, and *** denote statistical significance at the 10%, 5% and 1% levels.

	Heat. Oil	Heat.Oil	Nat. Gas	Nat. Gas	Crude Oil	Crude Oil	Copper	Copper
dNetCITshare	0.014	0.026	-0.278***	-0.305***	0.044	0.040	-0.002	-0.002
	(0.31)	(0.71)	(-2.26)	(-2.92)	(0.57)	(0.80)	(-0.56)	(-0.66)
dSpreadingshare	-0.038	-0.015	-0.034	0.047	0.127**	0.171***	-0.006	-0.003
	(-1.12)	(-0.48)	(-0.43)	(0.49)	(2.46)	(3.74)	(-0.70)	(-0.39)
dYieldSpread	-0.004		-0.007		0.006		-0.001	
	(-1.37)		(-1.16)		(1.05)		(-1.20)	
dStocks	0.004		-0.157***	-0.041**	-0.038	-0.073**	-0.003*	-0.003***
	(0.23)		(-2.87)	(-2.17)	(-0.85)	(-2.09)	(-1.98)	(-3.25)
dHedgePressure	0.007		0.054		0.005		0.007**	0.005***
	(0.37)		(1.14)		(0.12)		(3.23)	(2.97)
dTbillAdjusted	-0.002*		-0.005*	-0.005*	-0.006**	-0.004*	0.000	
	(-1.70)		(-1.87)	(-1.75)	(-2.04)	(-1.73)	(1.09)	
dTotalOI	0.006		-0.028		-0.013		0.001	
	(0.35)		(-0.75)		(-0.46)		(0.67)	
L.dSpread	-0.293*	-0.324***	0.285***	0.413***	0.054	0.104	-0.305**	-0.344**
	(-2.10)	(-2.67)	(3.16)	(6.79)	(0.28)	(0.54)	(-2.52)	(-2.49)
L.Spread	-0.613***	-0.500***	-0.905***	-0.766***	-0.452**	-0.412**	-0.620***	-0.528***
	(-3.18)	(-3.47)	(-8.09)	(-8.90)	(-2.38)	(-2.27)	(-3.69)	(-2.91)
L.NetCITshare	-0.043	-0.052**	0.075*	-0.011	-0.011	-0.024	-0.001	-0.003**
	(-1.63)	(-2.32)	(1.89)	(-0.64)	(-0.36)	(-1.24)	(-0.65)	(-2.38)
L.Spreadingshare	-0.033	-0.027	0.010	-0.000	-0.089**	0.009	0.004	0.005
	(-1.50)	(-1.46)	(0.13)	(-0.00)	(-2.16)	(0.73)	(0.52)	(0.85)
L.HedgePressure	0.008		0.036		-0.038		0.004**	
	(0.36)		(0.87)		(-0.90)		(2.17)	
L.Stocks	0.013	0.022***	-0.095***		0.033		-0.002***	-0.001***
	(1.38)	(4.13)	(-3.14)		(1.15)		(-3.22)	(-2.66)
L.YieldSpread	-0.005***	-0.005***	-0.003		-0.003*		-0.000	
	(-4.28)	(-4.27)	(-0.92)		(-1.69)		(-0.71)	
L.TotalOI	0.007		-0.006	-0.025	0.013		0.001	0.002*
	(1.05)		(-0.27)	(-1.48)	(0.57)		(1.24)	(1.75)
L.TbillAdjustedusted	-0.002**	-0.002**	-0.005**	-0.005**	-0.002		-0.000	
	(-2.65)	(-2.41)	(-2.47)	(-2.05)	(-1.41)		(-0.12)	
N	70	70	67	67	70	70	70	70
adj. R^2	0.383	0.385	0.619	0.625	0.318	0.265	0.542	0.518
F test for suppressed variables	F(7, 54) = 0.98 Prob > F = 0.46		F(6, 51) = 0.68 Prob > F = 0.66		F(8, 54) = 1.18 Prob > F = 0.33		F(6, 54) = 1.38 Prob > F = 0.24	
F-test (cointegration)	F(6, 61) = 4.43 Yes***		F(5, 57) = 45.75 Yes***		F(3, 62) = 2.89 No		F(5, 60) = 2.20 No	

Table 23. The impact of spreading positions on the spread of agricultural during the Goldman roll on the 8th business day of the month.

The ARDL model (4) is estimated for agricultural commodities. The dependent variable is the first difference of the spread. The monthly observations on the 8th business day are used. The description of other variables coincides with the previous models. First number in each cell is the OLS estimation of the coefficient for the corresponded variable. A number in brackets is the t-statistics. "N" is the number of observations. The symbols *, **, and *** denote statistical significance at the 10%, 5% and 1% levels.

	Cocoa	Cocoa	Coffee	Coffee	Corn	Corn	Soybeans	Soybeans	Sugar	Sugar	Wheat	Wheat
dNetCITshareroll	-0.010	-0.022	0.004	0.000	0.081	0.026	0.002	0.006	-0.138	-0.125*	-0.001	-0.028
	(-0.44)	(-1.16)	(0.43)	(0.01)	(1.58)	(0.57)	(0.06)	(0.42)	(-1.87)	(-1.96)	(-0.03)	(-1.05)
dSpreadingshareroll	0.004	0.004	0.001	0.008	-0.003	-0.008	-0.017	-0.010	0.091	0.097	0.037	0.011
	(0.21)	(0.22)	(0.09)	(0.81)	(-0.05)	(-0.11)	(-0.42)	(-0.30)	(0.57)	(0.66)	(1.12)	(0.39)
dYieldSpread	0.001		-0.000		0.003		0.004*	0.004	-0.003		0.002	
	(0.69)		(-0.56)		(1.30)		(1.83)	(1.55)	(-0.64)		(0.60)	
dStocks	-0.001		-0.018***	-0.013***	-0.007*		-0.003*	-0.001	0.006		0.001	
	(-0.25)		(-3.30)	(-3.65)	(-1.71)		(-1.85)	(-1.27)	(0.68)		(0.08)	
dHedgePressure	0.007		0.008***	0.008***	0.020	0.018	0.007		0.060**	0.046	0.017	
	(1.14)		(3.51)	(4.46)	(1.36)	(1.53)	(0.69)		(2.12)	(1.90)	(1.28)	
dTbillAdjusted	0.000		0.000		-0.000		0.002*	0.001*	-0.000		-0.000	
	(0.28)		(1.19)		(-0.40)		(1.96)	(1.82)	(-0.23)		(-0.62)	
dTotalOI	0.001		0.001		0.005		-0.006		-0.014		0.005	
	(0.36)		(0.82)		(0.45)		(-0.76)		(-0.82)		(0.59)	
L.dSpread	-0.295**	-0.389***	-0.269**	-0.313***	0.012	0.010	0.280	0.186*	-0.181	-0.134	-0.133	-0.310**
	(-2.34)	(-3.82)	(-2.52)	(-3.47)	(0.14)	(0.11)	(1.32)	(0.84)	(-1.47)	(-1.21)	(-1.01)	(-2.19)
L.Spread	-0.397***	-0.193**	-0.187**	-0.051**	-0.637***	-0.623***	-0.561***	-0.351**	-0.148	-0.137*	-0.472***	-0.125**
	(-2.89)	(-2.59)	(-2.04)	(-3.04)	(-9.43)	(-9.80)	(-2.91)	(-2.07)	(-1.58)	(-1.94)	(-3.18)	(-2.45)
L.NetCITshareroll	0.008	0.007	-0.002	-0.003	-0.004	-0.011	-0.008	-0.011	0.014	0.008	-0.009	-0.008*
	(0.76)	(0.72)	(-0.81)	(-1.38)	(-0.28)	(-0.84)	(-1.00)	(-1.49)	(0.58)	(0.44)	(-1.16)	(-1.81)
L.Spreadingshareroll	-0.010	-0.014	-0.001	0.001	0.003	0.017	0.014	0.013	-0.041	-0.047	0.007	0.003
	(-0.73)	(-1.11)	(-0.18)	(0.16)	(0.12)	(0.68)	(0.95)	(1.07)	(-0.85)	(-1.14)	(0.37)	(0.34)
L.HedgePressure	0.000		0.001		-0.006		0.014*		0.030	0.009	0.000	
	(0.12)		(0.53)		(-0.42)		(1.78)		(1.17)	(1.06)	(0.02)	
L.Stocks	-0.001		-0.001		-0.010***	-0.007***	-0.002		-0.006		0.001	
	(-0.39)		(-1.65)		(-4.30)	(-6.08)	(-1.83)		(-1.22)		(0.15)	
L.TotalOI	-0.001		0.000		0.004		-0.002		0.002		-0.003	
	(-0.38)		(0.07)		(0.98)		(-1.17)		(0.38)		(-0.76)	
L.YieldSpread	0.000		-0.000		0.001**	0.001**	0.001**	0.000	-0.000		-0.002**	
	(0.26)		(-0.73)		(2.09)	(2.92)	(2.17)	(1.29)	(-0.03)		(-2.04)	
L.TbillAdjusted	-0.000		-0.000		-0.000		0.000		0.001		-0.001	
	(-0.89)		(-0.43)		(-0.74)		(0.74)		(0.82)		(-1.37)	
N	81	82	81	82	78	78	81	81	78	81	81	82
adj. R^2	0.240	0.279	0.391	0.385	0.568	0.553	0.158	0.136	0.128	0.153	0.238	0.189

Table 24. The impact of spreading positions on the spread of commodities with the strong seasonality during the Goldman roll.

The ARDL model (4) is estimated for four commodities with the strong seasonality. The dependent variable is the first difference of the spread. The monthly observations on the 8th business day are used. All variables are described in the same way as for Table 15. First number in each cell is the OLS estimation of the coefficient for the corresponded variable. Monthly dummies are included where the number corresponds with the order of the month (m1 - January). The t-statistics of monthly dummy variables is suppressed. A number in brackets is the t-statistics. "N" is the number of observations. The results of the F-test for insignificant control variables are included under the "F-test of suppressed variables". The bottom line is the F-test for the presence of the cointegration of the level variables (lagged variables) in the right column model. The statistics compares with critical values of Pesaran, Table 8. If the value is higher the corresponding critical value, the absence of cointegration is rejected. In this case I insert "Yes" with the corresponding level of confidence. The symbols *, **, and *** denote statistical significance at the 10%, 5% and 1% levels.

	Gasoline	Gasoline	Lean Hogs	Lean Hogs	Live Cattle	Live Cattle	Cotton	Cotton
dNetCITshare	0.020	0.003	0.022	0.010	-0.119	-0.098	0.030	-0.008
	(0.33)	(0.06)	(0.27)	(0.13)	(-1.64)	(-1.51)	(0.56)	(-0.13)
dSpreadingshare	0.084	-0.029	0.058	0.068	0.047	0.049	0.036	0.032
	(0.90)	(-0.47)	(0.56)	(0.73)	(0.56)	(0.65)	(0.30)	(0.33)
dSpread	0.018*	0.018**	-0.002		-0.004		0.002	
	(2.33)	(2.43)	(-0.37)		(-1.23)		(0.53)	
dStocks	-0.071*		0.044		0.018		-0.008*	-0.005
	(-1.67)		(1.14)		(0.70)		(-2.53)	(-1.86)
dHedgePressure	0.013		0.012		-0.010		-0.005	
	(0.37)		(0.49)		(-0.45)		(-0.32)	
dTbillAdjusted	-0.003		0.002		-0.001		0.001	
	(-1.53)		(0.68)		(-0.75)		(0.89)	
dTotalOI	0.002		0.057**	0.040**	-0.031	-0.019	0.008	
	(0.07)		(2.43)	(2.02)	(-1.60)	(-1.24)	(0.52)	
L.dSpread	0.091	0.105	-0.051	-0.096	0.154	0.183	-0.107	-0.247
	(0.72)	(0.93)	(-0.39)	(-0.78)	(1.01)	(1.44)	(-0.76)	(-1.84)
L.Spread	-0.889***	-0.844***	-0.623***	-0.392***	-0.507***	-0.502***	-0.497***	-0.257***
	(-4.51)	(-4.19)	(-3.80)	(-3.32)	(-3.53)	(-4.58)	(-3.21)	(-2.66)
L.NetCITshare	-0.139**	-0.145***	-0.075	-0.098**	-0.060*	-0.066***	-0.030	-0.036
	(-2.47)	(-3.33)	(-1.53)	(-2.24)	(-2.00)	(-2.81)	(-0.93)	(-1.23)
L.Spreadingshare	-0.003	-0.143*	-0.000	-0.010	0.054	0.041	0.107	0.047
	(-0.03)	(-1.67)	(-0.00)	(-0.14)	(0.90)	(0.87)	(1.08)	(0.48)
L.HedgePressure	-0.040		0.038*		0.020		-0.005	
	(-0.82)		(1.84)		(1.04)		(-0.51)	
L.Stocks	0.064**	0.052***	0.002		0.004		-0.003	
	(2.46)	(4.43)	(0.13)		(0.21)		(-1.43)	
L.TotalOI	-0.012		0.014		-0.017		0.015	
	(-0.61)		(0.70)		(-1.43)		(1.40)	
L.TbillAdjusted	0.002		0.001		-0.000		0.000	
	(1.44)		(0.52)		(-0.12)		(0.58)	
L.Spread			0.001		0.000		0.003**	
			(0.39)		(0.53)		(2.23)	
m1	-0.044***	-0.045***	0.027	0.035	0.012	0.016	-0.006	-0.008
m2	0.017	0.017	-0.026***	-0.029***	0.024***	0.023***	-0.012	-0.004
m3	-0.004	-0.003	-0.000	0.017	0.023	0.025	-0.006	-0.005
m4	0.002	0.005	0.028***	0.034***	-0.002	-0.005	-0.004	0.005
m5	0.008	0.011*	0.048	0.053	0.022	0.024	-0.006	-0.004
m6	0.013**	0.016**	0.044	0.045	-0.015***	-0.016***	-0.002	0.005
m7	0.064***	0.068***	0.095**	0.092**	0.016	0.020	-0.009	-0.012*
m8	0.010	0.011	0.020	-0.004	0.011*	0.011**	-0.008	-0.010*
m9	0.020	0.028**	0.056	0.050	0.015	0.020	-0.012*	-0.014**
m10	0.009	0.013*	-0.020	-0.035***	-0.000	-0.000	-0.010	-0.012
m11	-0.001	0.004	0.018	0.025	0.016	0.019	-0.014	-0.003
constant	-0.013	-0.004	-0.033	-0.006	0.005	-0.005	-0.010	0.007
	(-0.58)	(-0.30)	(-1.28)	(-0.86)	(0.27)	(-1.92)	(-0.89)	(1.70)
N	70	70	80	80	81	82	81	82
adj. R^2	0.909	0.905	0.840	0.836	0.741	0.750	0.182	0.207

Reference

1. Acharya, V., Lochstoer, L., & Ramadorai, T. (2013). Limits to Arbitrage and Hedging: Evidence from Commodity Market. *Journal of Financial Economics* (109), 441-465.

2. Akram, Q. F. (2009). Commodity Prices, Interest Rates and the Dollar. *Energy Economics* (31(6)), 838-851.

3. Anson, M. (1998). Spot Returns, Roll Yield, and Diversification with Commodity Futures. *Journal of Alternative Investment , Winter*, 16-32.

4. Bailey, W., & Chan, K. C. (1993). Macroeconomic Influences and the Variability of the Commodity Futures Basis. *Journal of Finance , 48* (2), 555-573.

5. Baker, S. D. (2012). The Financialization of Storable Commodities. *Working Paper* .

6. Basak, S., & Pavlova, A. (2013). A Model of Financialization of Commodities. *Working Paper* .

7. Basak, S., & Pavlova, A. (2013). Asset prices and institutional investors. *American Economic Review* (forthcoming).

8. Basu, D., & Miffre, J. (2013). Capturing the Risk Premium of Commodity Futures:The Role of Hedging Pressure. *Working Paper* .

9. Bichetti, D., & Maystre, N. (2012). The Syncronized and Long-lasting Structural Change on Commodity Markets: Evidence from High Frequency Data. *Working Paper* (No.37485).

10. Bowsher, C., & Meeks, R. (2006). The Impossibility of Stationary Yield Spreads and I(1) Yields under the Expectations Theory of the Term Structure. *Working Paper* .

11. Brunetti, C., & Reiffen, D. (2011). *Commodity Index Trading and Hedging Costs.* Division of Research & Statistics and Monetary Affairs. Washington D.C.: Federal Reserve Board.

12. Brunner, A. D. (2002). El Nino and world primary commodity prices: warm water or hot air? *Review of Economics and statistics , 84* (1), 176-283.

13. Buyuksahin, B., & Robe, M. (2011). Does "Paper Oil" Matter? Energy Markets' Financialization and Equity-commodity Co-movements. *Working Paper* .

14. CFTC. (2009). *Staff Report on Commodity Swap Dealers and Index Traders with Commission Recommendations.* U.S. Commodity Futures Trading Commission.

15. Daskalaki, C., & Skiadopoulos, G. (2011). Should Investors Include Commodities in Their Portfolios After All? New Evidence. *Journal of Banking & Finance , 35* (10), 2606-2626.

16. de Roon, F., Nijman, T. E., & Veld, C. (2000). Hedging Pressure Effects in Futures Markets. *Journal of Finance , 55* (3), 1437-1456.

17. Deaton, A., & Laroque, G. (1992). On the Behavior of Commodity Prices. *Review of Economic Studies , 59*, 1-23.

18. Dempster, M. A., Medova, E., & Tang, K. (2012). Determinants of Oil Futures Prices and Convenience Yields. *Quantitative Finance , 12* (12), 1795-1809.

19. Dincerler, C., Khokher, Z., & Simin, T. (2006). An Empirical Analysis of Commodity Convenience Yields. *Working Paper* .

20. Domanski, D., & Heath, A. (2007, March). Financial Investors and Commodity Markets. *Bank of International Settlements Quarterly Review* , 53-67.

21. Dong, F., Du, X., & Gould, B. W. (2011). Milk Price Volatility and its Determinants. *2011 AAEA & NAREA Joint Annual Meeting.*

22. Erb, C. B., & Harvey, C. R. (2006). The Strategic and Tactical Value of Commodity Futures. *Financial Analysts Journal , 62*, 69-97.

23. Estrella, A., & Mishkin, F. (1998). Predicting U.S. Recessions: Financial Variables as Leading Indicators. *Review of Economics and Statistics , 80*, 45-61.

24. Fama, E. F., & French, K. R. (1988). Business Cycles and the Behavior of Metals Prices. *Journal of Finance , 43* (5), 1075-1093.

25. Feldman, B., & Till, H. (2006). Backwardation and Commodity Futures Performance: Evidence from Evolving Agricultural Markets. *Journal of Alternative Investment , Winter*, 24-39.

26. Frankel, J. A. (2006). The Effect of Monetary Policy on Real Commodity Prices. *Working Paper* .

27. Frenk, D., & Turbeville, W. (2011). Commodity Index Traders and the Boom/Bust Cycle in Commodities Prices. *Working Paper* .

28. Friedman, M. (1953). The Case for Flexible Exchange Rates. In M. Friedman, *Essays in Positive Economics.* Chicago,IL: University of Chicago Press.

29. Gilbert, C. L. (2009). Speculative Influences on Commodity Futures Prices, 2006-2008. *Working Paper* .

30. Gorton, G., & Rouwenhorst, K. G. (2006). Facts and Fantasies about Commodity Futures. *Financial Analysts Journal , 62* (2), pp. 47- 68.

31. Gorton, G., Hayashi, F., & Rouwenhorst, G. (2007). The Fundamentals of Commodity Futures Returns. *Working Paper* .

32. Granger, C. W., & Newbold, P. (1974). Spurious Regressions in Econometrics. *Journal of Econometrics , 2* (2), 111-120.

33. Greenwood, R. (2005). Short- and Long-term Demand Curves for Stocks: Theory and Evidence on the Dynamics of Arbitrage. *Journal of Financial Economics , 75*, 607-649.

34. Grosche, S. (2012). Limitations of Granger Causality Analysis to Assess the Price Effects from the Financialization of Agricultural Commodity Markets under Bounded Rationality. *Working Paper* .

35. Hamilton, J. D. (2011). Nonlinearities and the Macroeconomic effects of Oil Prices. *Macroeconomics Dynamics , 15* (S3), 364-378.

36. Hamilton, J. D. (2009). Understanding Crude Oil Prices. *Energy Journal , 20* (2), 179-206.

37. Hong, H., & Yogo, M. (2009). Digging into Commodities. *Working Paper* .

38. Hong, H., & Yogo, M. (2012). What does Futures Market Interest Tell Us about the Macroeconomy and Asset Prices? *Journal of Financial Economics , 105* (3), 473-490.

39. Irwin, S. H., & Sanders, D. R. (2011). Index Funds, Financialization, and Commodity Futures Markets. *Applied Economic Perspectives and Policy , 33*, 1-31.

40. Irwin, S. H., & Sanders, D. R. (2013). Measuring Index Investment in Commodity Futures Markets. *Energy Journal , forthcoming*.

41. Irwin, S. H., & Sanders, D. R. (2011). Testing the Masters Hypothesis in Commodity Futures Markets. *Energy Economics , 34*, 256-269.

42. Irwin, S. H., & Sanders, D. R. (2010). The Impact of Index and Swap Funds on Commodity Futures Markets: Preliminary Results. *Working Paper* .

43. Janzen, J. P., Smith, A. D., & Carter, C. A. (2012). Commodity Price Comovement: The Case of Cotton. *Working Paper* .

44. Kaldor, N. (1939). Speculation and Economic Stability. *Review Economic Studies , 7*.

45. Keynes, J. M. (1930). *A Treatise on Money, vol. 2.* London: Macmillan & Co. Ltd.

46. Kilian, L. (2009). Not All Oil Price Shocks are Alike: Disentangling Demand and Supply Shocks in the Crude Oil Market. *American Economic Review , 99* (3), 1053-1069.

47. Litzenberger, R., & Rabinowitz, N. (1995). Backwardation in Oil Futures Markets: Theory and Empirical Evidence. *Journal of Finance , 50* (5), 1517-1545.

48. Long, J. B., Shleifer, A., Summers, L. H., & Waldmann, R. J. (1990). Positive Feedback Investment Strategies and Destabilizing Rational Speculation. *Journal of Finance , 45* (2), 379-395.

49. Masters, M. W., & White, A. K. (2008). The Accidental Hunt Brothers: How Institutional Investors are Driving up Food and Energy Prices. *Working Paper* .

50. Mou, Y. (2011). Limits to Arbitrage and Commodity Index Investment: Front-Running the Goldman Roll. *Working Paper* .

51. Mu, X. (2007). Weather, storage, and natural gas price dynamics: Fundamentals and volatility. *Energy Economics , 29* (1), 46-63.

52. Norrish, K., & Croft, H. (2013). The Commodity Refiner from an Age of Shortage to an Era of Enough. *Barclays Capital Research* .

53. Pesaran, M. H., & Shin, Y. (1999). An Autoregressive Distributed Lag Modelling Approach to Cointegration Analysis. In S. Stom, *Econometrics and Economic Theory in the Twentieth Century.* New York, NY: Cambridge University Press.

54. Pesaran, M. H., Shin, Y., & Smith, R. J. (2001). Bounds Testing Approaches to the Analysis of Level Relationships. *Journal of Applied econometrics , 16* (3), 289-326.

55. Petzel, T. (2009). Testimony before the CFTC. *Working Paper* .

56. Roll, R. (1984). Orange Juice and Weather. *American Economic Review , 74* (5), 861-880.

57. Rossi, C. V. (2011). Analysis of CFTC Proposed Position Limits on Commodity Index Fund Trading. *Working Paper* .

58. Silvennoinen, A., & Thorp, S. (2012). Financialization, Crisis and Commodity Correlation Dynamics. *Journal of International Financial Markets, Institutions & Money , 24*, 42-65.

59. Singleton, K. (2011). Investor Flows and the 2008 Boom/Bust in Oil Prices. *Working Paper* .

60. Sockin, M., & Xiong, W. (2012). Feedback Effects of Commodity Futures Prices. *Working Paper* .

61. Stoll, H., & Whaley, R. (2010). Commodity Index Investing and Commodity Futures Prices. *Journal of Applied Finance , 20*, 7-46.

62. Tang, K., & Xiong, W. (2010). Index Investments and Financialization of Commodities. *Working Paper* .

63. Working, H. (1949). The Theory of Price of Storage. *American Economic Review , 29.*