Centre for Climate Change Economics and Policy

An ESRC Research Centre

Grantham Research Institute on Climate Change and the Environment

Comovement and the financialization of commodities

Matteo Bonato and Luca Taschini

November 2015

Centre for Climate Change Economics and Policy
Working Paper No. 240

Grantham Research Institute on Climate Change and the Environment
Working Paper No. 215

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL

The Grantham Foundation for the Protection of the Environment

University of Leeds

LSE The London School of Economics and Political Science

The Centre for Climate Change Economics and Policy (CCCEP) was established by the University of Leeds and the London School of Economics and Political Science in 2008 to advance public and private action on climate change through innovative, rigorous research. The Centre is funded by the UK Economic and Social Research Council. Its second phase started in 2013 and there are five integrated research themes:

1. Understanding green growth and climate-compatible development
2. Advancing climate finance and investment
3. Evaluating the performance of climate policies
4. Managing climate risks and uncertainties and strengthening climate services
5. Enabling rapid transitions in mitigation and adaptation

More information about the Centre for Climate Change Economics and Policy can be found at: http://www.cccep.ac.uk.

The Grantham Research Institute on Climate Change and the Environment was established by the London School of Economics and Political Science in 2008 to bring together international expertise on economics, finance, geography, the environment, international development and political economy to create a world-leading centre for policy-relevant research and training. The Institute is funded by the Grantham Foundation for the Protection of the Environment and the Global Green Growth Institute. It has nine research programmes:

1. Adaptation and development
2. Carbon trading and finance
3. Ecosystems, resources and the natural environment
4. Energy, technology and trade
5. Future generations and social justice
6. Growth and the economy
7. International environmental negotiations
8. Modelling and decision making
9. Private sector adaptation, risk and insurance

More information about the Grantham Research Institute on Climate Change and the Environment can be found at: http://www.lse.ac.uk/grantham.

Comovement and the Financialization of Commodities*

Matteo Bonato[a †] Luca Taschini[b, c, d‡]
[a] University of Johannesburg
[b] Grantham Research Institute on Climate Change and the Environment
[c] ESRC Centre for Climate Change Economics and Policy
[d] London School of Economics and Political Science

November 2015

Abstract

Commodity markets have experienced a surge in participation by financial investors over the past decade. This notable uptick highlighted the potential effect of financialization as one of the causes of increased comovement of commodities returns. Building on the theory of comovement proposed by Barberis et al. (2005), this paper investigates drivers of the observed increase in comovement prices of commodity futures in the United States and provides new empirical evidences. This empirical analysis allows us to distinguish between two views of return comovement: the traditional view, in which return comovement is explained by comovements in news about fundamentals value, and the alternative view, in which excess comovement is attributed to traders' sentiment. We examine differences in the dependence structure between index and off-index commodities, and three major commodities indexes. Non-energy commodities present in the index exhibit an increase in comovement whereas those commodities off the index do not. These results are supported by the analysis of high-frequency returns dynamics by means of the so called realised beta. We therefore provide new evidence supporting the friction or sentiment based view of commodity comovement. We interpret this to be additional evidence of financialization as the driver of the observed increase in commodity return comovement.

Keywords: Commodities; Comovement; Financialization; Indexing; Realized Beta.

*Part of this paper was written while Matteo Bonato was visiting the University of Pretoria and the University of Zimbabwe, whose hospitality is greatly acknowledged. Part of Taschini's research was supported by the ESRC Centre for Climate Change Economics and Policy, which is funded by the UK Economic and Social Research Council. The authors would like to thank René Carmona, Jennifer Mills, Mindy Mallory, Anna Pavlova, Michel Robe, Wei Xiong, and other participants of the Workshop on Commodity Markets and their Financialization held at IPAM, UCLA. The usual disclaimers apply.

†Address: Department of Economics and Econometrics, University of Johannesburg, Auckland Park, South Africa. E-mail: `mbonato@uj.ac.za`

‡Address: The Grantham Research Institute on Climate Change and the Environment, London School of Economics and Political Science, UK. E-mail: `l.taschini1@lse.ac.uk`.

1 Introduction

Commodity futures markets have experienced a surge in participation by financial investors over the past decade, as investors turn to commodities as a means to diversify their portfolios while maintaining returns comparable to or better than the major stock indices (Cheng and Xiong (2014)). Institutional investors and private wealth managers have long sought alternative asset classes to hedge against inflation and systemic stock market fluctuations. Examples of these alternative asset classes include traditional investments such as real estate, private equity, and hedge fund investments, as well as more niche investments in art and other collectibles (Ankrim and Hensel (1993), Edwards and Liew (1999), Campbell (2008), Dimson and Spaenjers (2011)). These asset classes are sought out given their seeming low or even negative correlation with equity and bond markets. Before the early 2000s, commodity markets also exhibited these characteristics; there is evidence that commodity markets were at least partially segmented from both outside financial markets and each other. Working with historic price data from 1982-2004, Erb and Harvey (2006) demonstrated that commodities in different sectors had little price correlation with each other, while Gorton and Rouwenhorst (2006) showed that commodity futures returns were actually negatively correlated with equity and bond returns over the 1959-2004 period.

The value of commodities investment is two-fold: commercial participants in the commodities market (those physically involved with the production and consumption of commodities) use commodity derivatives markets to hedge against future price fluctuations, while non-commercial market participants utilize commodities solely to improve or diversify their portfolios (Cheng and Xiong (2014)). The attractiveness of commodities as a hedging mechanism for non-commercial market participants is augmented by the generally high returns provided by investing in commodity futures contracts; Gorton and Rouwenhorst (2006) show that both the average annual return and return volatility from commodities is comparable to that from the S&P 500 over their 1959-2004 dataset. The number of these non-commercial market participants seeking portfolio diversification expanded rapidly in the early 2000s following the collapse of the equity market in 2000. Many financial institutions and indexers promoted commodities as a new asset class, and as a result, billions of investment dollars flowed into commodity markets from financial institutions, insurance companies, pension funds, foundations, hedge funds, and wealthy individuals. Particularly striking was the increase in popularity of commodity indices (including the Standard & Poor's Goldman Sachs commodity index, the Dow Jones-UBS Commodity Index and the Thomson-Reuters Commodity Index), which allowed for ready access to the commodities market by a wide range of financial players (Irwin and Sanders (2012), Masters (2008), and Adams and Gluck (2015)).

Perhaps unsurprisingly, such large index investment flow had a measurable impact on the behaviour of commodities markets. Analysts studying commodity price fluctuations found increasing correlations between the returns on different classes of commodities post-2004, suggesting that commodity markets were becoming less segmented from each other (Tang and Xiong (2012), Hamilton and Wu (2014), Gilbert (2010), Daskalakis et al. (2014)). This, in turn, led to speculation about the negative effects of the "financialization" of commodity markets, here defined to mean the excessive role of financial markets, financial actors, and financial institutions in commodities markets (Cheng and Xiong (2014) and Cheng et al. (2014) offer empirical investigations of financialization; Basak and Pavlova (2014) offer a theoretical treatment of financialization of commodities). In particular, a synchronized boom and bust cycle in 2007 and 2008 was observed in several commodities across the energy, metal, and agricultural sectors. Given that commodities are often core inputs to production processes or consumptive goods (food price volatility was a major concern), the potential negative impact of this increased financialization became a serious public policy concern (Masters, 2008). However, the underlying cause of the observed increase in commodity price comovement remains a question of considerable debate. While some argue that the increased price correlation is due to this commodity financialization phenomenon and the impact of increased speculative activities (e.g. Tang and Xiong (2012), Cheng et al. (2014), Hamilton and Wu (2014), and Adams and Gluck (2015)), others maintain that the concurrent movements can be explained by fundamental supply and demand (Krugman (2008), Stoll and Whaley (2010), Irwin et al. (2009) and Fattouh et al. (2013)). Still others have argued for a combination of these explanations, where increased speculative activity acts as an enhancer, but not a fundamental driver, of commodity and equity return comovements (Bruno et al. (2015)).

This paper provides further evidence for the commodity financialization view through an analysis of the drivers of the observed commodity price comovement, based upon the theory of comovement proposed by Barberis, Shleifer and Wurgler (2005). In particular, we investigate the difference between index and off-index commodities considering 25 commodities (no pooling), with the aim of complementing the work by Tang and Xiong (2012). The difference between index and off-index commodities is the focus of this analysis because index investors tend to move in and out of all commodities in a given index at the same time (as these investors tend to focus on their portfolio allocation between commodities in general and other asset classes such as equities and bonds). Barberis and Shleifer's (2003) findings on 'style investing' would suggest that changes in funds allocation could cause the prices of commodities included in the fund or index to move together. Thus one would expect price comovements of commodities included in an index to be greater than comovements of off-index commodities. Barberis et al. (2005) observed this in the equity market, finding that a stock's listing on the S&P 500 could significantly increase its return correlation with the index.

Motivated by these studies, we test these alternative theories of comovement and exploit changes in the return comovement of index and off-index commodities futures in the Unites States market with three of the most popular commodity indices in order to provide new evidence of the financialization process. Our contribution to the literature on the commodity financialization is two fold. First, we investigate the drivers of the commodity financialization process testing the theories of investment behaviour suggested by Barberis et al. (2005) but using a different argumentation. Barberis et al.

(2005) propose an alternative driver for the comovement observed amongst stocks included in the S&P 500. Traditionally, comovement in prices is interpreted to reflect comovement in fundamentals value. Barberis et al. (2005) propose an alternative theory in which comovement arises when investors group assets in categories and trade at the category, rather than individual asset level, or when traders choose to trade only a subset of the available assets (referred to for the remainder of this paper as the 'category/habitat' or 'sentiment' view).

We argue that the financialization of commodities and the availability of new, commodity-related investment instruments had an effect on commodity returns similar to that of a stock's inclusion in a major index: comovement with the in-the-index commodities increases and comovement with the off-the-index commodities decreases. We test this hypothesis through a series of univariate and bivariate regressions, examining two time periods: 1998-2004 (pre-financialization) and 2005-2011 (post-financialization). The choice of 2005 as the breakpoint is innocuous; it was chosen as it is generally agreed in the literature that the increased price comovements between various commodities started after 2004 (Tang and Xiong (2012)).[1] The aim is to search for changes in the return comovements between the three main commodity indexes (Standard & Poor's GSCI, the Dow Jones-UBS Commodity Index, and the Thomson-Reuters Commodity Index) and commodities included and not included in the indexes. An increase in price comovement between index commodities and the index indicates the existence of a category/habitat view; we interpret this to be evidence of financialization.

In the univariate regressions, we observe a significant increase in comovement between index non-energy commodities and index starting in 2005. For off-index commodities, either no change or a significant decrease in comovement is observed. Thus after 2005, two distinct groups of commodities emerged - financialized commodities (those included in the major indices) and non-financialized commodities (off-index commodities). These results were corroborated and more finely tuned through the bivariate regressions: again index non-energy commodities exhibit an increase in comovement with the index while off-index commodities do not exhibit increased comovement with the index. While we interpret these results as evidence of increased financialization post-2005, alternative explanations for the regression findings include non-trading effects (i.e. that the comovement results could have a spurious upward bias due to the greater liquidity and increased trading activity of index commodities) and characteristic effects (i.e. that all commodities included in the indices share a common characteristic and that characteristic increases with time). We explicitly test for both of these effects and find that their impact is likely to be minimal. We also re-run the regressions for an earlier (pre-financialization) period and find no evidence of an increase in comovement between index or non-index commodities and the indices, which provides further support for the finacialization interpretation.

[1] It should be noted that some authors argue that the shift in correlations between commodities and equities and between commodities did not begin until 2008 (e.g. Adams and Gluck (2015) and Buyuksahin and Robe (2014)). While we maintain that evidence of increased comovement can be seen before the financial crisis, changing the cut-off date to 2008 does not materially change the results of our analysis. Data on univariate regressions run for the pre-/post-2008 periods is available upon request.

Our second major contribution to the literature is that, relying on the availability of high-frequency data, we can go one step beyond daily price regressions and calculate the so called realised beta between each commodity and the index. These realised beta calculations are used to further test for the presence of financialization in commodity markets post-2005. We observe a statistically significant increase in realized betas of index non-energy commodities and index, and a smaller or non significant increase in realized betas of off-index non-energy commodities and index. These results support the regressions on daily returns and indicate that on average, realized betas of index commodities increase after 2004. Moreover, this increase in realised betas tells us that the increase in return comovement can be seen not only at a daily frequency, but even at intra-day levels. This provides further evidence that the increase in comovement observed is not related to changes in fundamentals and instead represents an increase in commodity financialization.

It is also worth noting that our empirical findings of increased comovement between index commodities and the index are in line with the predictions of a fully-rational theoretical model developed by Basak and Pavlova (2014). Basak and Pavlova (2014, 2013) develop a theoretical model that features institutional investors (who care about their performance relative to a commodity index) alongside traditional commodity futures market participants and demonstrate that this addition of institutional investors can lead to an increase in correlation amongst commodity futures. Although additional empirical work is necessary to explicitly test the predictions of this alternate model of commodity comovement, it is interesting to note that the empirical behaviour we observe could potentially also be explained by a fully-rational model of investor behaviour, in addition to the behavioural model analysed here.

The first part of Section 2 is devoted to the discussion of the theory of comovement and the description of the testable predictions. Results of the univariate and bivariate regressions are dealt with in the second part of Section 2, while the final part of this section considers alternative explanations for the results of the regressions and provides robustness checks. Section 3 concludes the empirical analysis presenting the results of the test for the presence of financialization in commodity using realised betas. Section 4 concludes.

2 The Theory of Comovement and the Empirical Tests

The primary goal of this work is to explore the drivers of the observed increase in commodity return comovement pre- and post-2005, drawing from the theory of comovement proposed by Barberis et al. (2005). In this section, we introduce Barberis, Shleifer and Wurgler's take on comovement and introduce the testable predictions that will form the basis of our analysis.

The backbone of the analysis is a comparison between the return dependence structure for commodities included in major commodity indices (hereafter index commodities) and those not included in these indices (hereafter off-index commodities). We explore differences between index and off-index commodities because of the rise in commodity index fund (as well as ETF) investing following the 2000 crash in the equity market. The potential issue with increased cash flows to index funds is that index investors tend to move in and out of all commodities in a given index at the same time (as these investors tend to focus on their portfolio allocation between commodities in general and other asset classes such as equities and bonds). Changes in funds allocation can thus cause commodities in the fund to move together (Barberis and Shleifer (2003)), meaning that one would expect price comovements of commodities included in an index to be greater than comovements of off-index commodities.

Barberis et al. (2005) formalize this potential driver of return comovement. According to traditional theory, comovement in prices should reflect comovement in fundamental value (the asset's discounted future cash flows for stocks and the supply-demand balance for commodities). Barberis, Shleifer, and Wurgler postulate that in economies with less than perfectly rational investors, comovement can be delinked from fundamental values when investors group assets in categories and allocate funds at category level rather then at individual asset level (the category view) or when traders choose to trade only a subset of all available asset (the habitat view). They test this by investigating whether the inclusion of a stock to the S&P 500 (index) leads to a shift in the dependence structure of returns.

In this analysis, we reproduce Barberis, Shleifer, and Wurgler's reduced-form model of comovement. We consider an economy with $2n$ commodities. Each commodity i represents a commodity futures contract. These contracts are never held until physical delivery but each time before maturity, are rolled over to the next available liquid contract. Suppose that, to simplify their decision-making, investors group the $2n$ commodities into two categories, IND and OFF, and then allocate funds at the level of these categories rather than at the individual commodity level. It may be helpful to think of IND and OFF as index and off-index commodities, respectively. Suppose now that investors channel funds in and out of the categories depending on their sentiment.

A simple representation for commodity returns is then:

$$\Delta P_{i,t} = \epsilon_{i,t} + \Delta u_{IND,t}, \quad i \in IND, \tag{1}$$

$$\Delta P_{j,t} = \epsilon_{j,t} + \Delta u_{OFF,t}, \quad j \in OFF. \tag{2}$$

where $u_{IND,t}$ and $u_{OFF,t}$ are jointly normally distributed with means zero and same variance σ_u^2. $\Delta P_{i,t} \equiv P_{i,t} - P_{i,t-1}$ where $P_{i,t}$ are log-prices; and $\Delta\epsilon_t = (\Delta\epsilon_{1,t}, \ldots, \Delta\epsilon_{2n,t})' \sim N(0, \Sigma)$, i.i.d. over time.

Here $\epsilon_{i,t}$ is an idiosyncratic term denoting the flow of news for commodity i related to fundamentals (demand and supply); $u_{IND,t}$ can be thought of as time t index trader sentiment about the commodities in category IND, e.g. the commodity index. Since the index traders allocate funds by category, this sentiment level is the same for all commodities in category IND. Equation (1) says that the return on a commodity in category IND is affected not only by news ϵ but also by the change in sentiment about IND, $\Delta u_{IND,t}$; when index traders become more bullish about index commodities, the commodities in the index go up in price. The same applies for commodity in category OFF.

To uncover evidence on the alternative theory of comovement (category and habitat) we look for testable predictions. Let's start with Prediction 1 in Barberis et al. (2005), which describes what happens when a commodity is added to an index (i.e. moves from OFF to IND) in a univariate regression.

Prediction 1. *Suppose that the commodity j, previously a member of OFF, is now classified into IND. Then for a large number of commodities n, the probability limit of the OLS estimate of β_j in the univariate regression*

$$\Delta P_{j,t} = \alpha_j + \beta_j \Delta P_{IND,t} + \nu_{j,t}, \tag{3}$$

where

$$\Delta P_{IND,t} = \frac{1}{n} \sum_{l \in IND} \Delta P_{l,t},$$

as well as the probability limit of the R^2 of this regression, increase after reclassification.

In other words, the presence of commodity j in category IND increases the covariance of its return with the return on the category. The intuition behind this is relatively straightforward: when commodity j enters category IND, it is hit by index traders' flows of funds into and out of the category, which increases the covariance of its return with the return of category IND ($\Delta P_{IND,t}$) and hence also its beta loading on that return.

We now introduce Prediction 2 from Barberis et al. (2005), which describes what happens when a commodity is added to an index in a bivariate regression:

Prediction 2. *Suppose that the commodity j, previously a member of OFF, is now classified into IND. Then for a large number of commodities n, the probability limit of the OLS estimate of $\beta_{j,IND}$ in the bivariate regression*

$$\Delta P_{j,t} = \alpha_j + \beta_{j,IND} \Delta P_{IND,t} + \beta_{j,OFF} \Delta P_{OFF,t} + nu_{j,t} \tag{4}$$

rises after reclassification while the probability limit of the OLS estimate of $\beta_{j,OFF}$ falls.

This simply says that when a commodity enters the IND category, it becomes more sensitive to the IND category sentiment shock, $\Delta u_{IND,t}$ from Equation (1). This second prediction is important because the independent variable in the Prediction 1 regression, $\Delta P_{IND,t}$, is not a clean measure of this sentiment shock: as Equation (1) shows, a substantial part of its variation comes from news related to the commodity fundamentals, ϵ_t. The $\Delta P_{OFF,t}$ variable in the regression of Prediction 2 can be thought of as a control for such news, making the coefficient on $\Delta P_{IND,t}$ a cleaner measure of sensitivity to $\Delta u_{IND,t}$.

Barberis et al. (2005) use the bivariate regression test described by Prediction 2 to test whether the inclusion of a stock to the S&P 500 (index) leads to a shift in the correlation structure of returns. The alternative theory implies that, in a bivariate regression of the stock's return on both the S&P 500 index and non-S&P "rest of the market" index, when a stock is added to the index, the S&P beta should increase and the non-S&P (rest of the world) beta should decrease. Conversely, when a stock is removed from the index, the S&P beta should decrease and the non-S&P (rest of the world) beta should increase. Moreover, given the growing importance of the S&P 500 as both a category and a habitat, the changes in comovement should be stronger in more recent data. Barberis and colleagues find that these predictions hold true for stock returns from 1976-2000.

In this study, we run an analogous analysis for the commodities market, examining the difference between returns of commodities included in the three major commodity indices (Standard & Poor's GSCI (S&P GSCI), the Dow Jones-UBS Commodity Index (DJ-UBS) and the Thomson-Reuters Commodity Index (Thomson-Reuters CI)) and those not included in the indices. To test the model predictions on a group of commodities, three conditions are required:

i the group must be a natural category or preferred habitat for investors;

ii since the predictions focus on "reclassification", there must be clear and identifiable changes in commodity group membership; and

iii to control for fundamentals-based comovement, a commodity's inclusion or removal from the index should not change investors' perception of the correlation of the commodity's fundamental value with the fundamental values of other commodities in the index.

Stocks in the S&P 500 index satisfy all three conditions. For the case of commodity indices, the situation is slightly more complex and needs to be discussed further. We now examine each one of the three requirements to assess to what degree they are satisfied and discuss associated impacts for our analysis:

(i) Commodity indices must be a natural category or preferred habitat for investors. Commodity indices have indeed become a preferred category/habitat for investment due to the changing nature of commodity futures market participation. As alluded to previously, over the last decade there has been a large inflow of investment capital from investors looking to commodities for diversification purposes as part of a broader portfolio strategy as opposed to commercial commodity market participants looking to hedge spot-price risk (Cheng and Xiong (2014) and Adams and Gluck (2015)). These investors tend to treat commodity futures as an asset class just like stocks and bonds, and often invest in financial instruments linked to the large commodity indices. Analogous to the S&P 500 index, the considerable popularity of the products linked to commodity indices suggests that the indices are a preferred habitat for some investors (i.e. a subset of securities that the investors choose to trade exclusively) and a natural category for many more (i.e. investors allocate funds to the index rather than at the individual asset level).

(ii) There must be clear and identifiable changes in commodity group membership. For stocks in the S&P 500, there are clearly identifiable entry and exit dates and thus identifiable changes in group membership. Moreover, there is enough turnover in membership (30 changes in a typical year) to perform a traditional event study. While in principle commodities indices also possess clear and identifiable turnover in membership, the trading history and membership turnover does not allow us to perform a proper event study as in Barberis et al. (2005). Since the creation of the major commodities indices, only a few commodities have been added, while the main core of the indices have remained unchanged.[2] Thus there is not a robust enough dataset of commodity entry and exit from the indices to perform an event study based on individual commodity entry and exit dates. However, a turning point in investment activity can still be identified. Although commodity indices have existed since the early 1990s,[3] the inflow of investment funds increased dramatically only after 2004 (Tang and Xiong (2012), Basak and Pavlova (2014)). We therefore consider 2005 as turning point for the financialization of commodities as it marks the increase of inflow capital, which triggered the increase in commodity comovement. Thus a single time point is used, pre- and post-2005, in line with previous analysis (Tang and Xiong (2012)).

(iii) A commodity's inclusion or removal from the index should not change investors' perception of the correlation of the commodity's fundamental value with the fundamental values of other commodities in the index. For the S&P 500, the stated goal of the index is to be representative of the broader U.S. economy, not to signal a view about future cash flows. Analogously, commodities indices aim to represent the commodities futures market as a whole and thus inclusion of a commodity in an index should not signal anything about the commodity's fundamentals.

2.1 Data

We consider a total of 25 commodity futures traded in the United States market.[4] Table 1 presents these commodities along with their sector, ticker, the exchange in which they are traded, and their weight in the three major indices considered here: S&P GSCI, the DJ-UBS commodity index, and the Thomson-Reuters commodity index. The S&P GSCI features 18 US commodities, and offers the least diversified investment exposure of the big three. It is heavily weighted towards energy, mostly oil. The DJ-UBS commodity index includes 16 US commodities with specific minimum and maximum weightings; a commodity should have at least a 2% weighting and a maximum allocation of 25%. The Thomson-Reuters was first calculated by Commodity Research Bureau in 1957 and of the three is unique in that all 17 US commodities included are weighted equally at 5.88% per component.

Futures data is in continuous format. Close to expiration of a contract, the position is rolled over to the next available contract, provided that activity has increased. Our dataset spans the period from April 9, 1998 to March 24, 2011 for a total of 3,222 trading days, using 1-minute frequency data.[5] In order to guarantee that our results are based on overlapping time periods, we only consider trades

[2] For instance, Soybean Meal was included in DJ-UBS starting from January 2013. See DJ-UBS index weblink.

[3] S&P GSCI has existed since 1991 and DJ-UBS has existed since 1998.

[4] Since Kansas Wheat has positive weight for S&P GSCI but it is not available in our sample, we allocate its weight to the Chicago Wheat.

[5] The minute frequency data is constructed using the last price that occurs in the minute using data provided by Disk Trading.

that occurred between 10.30 and 14.00 NYT.[6] Daily returns close-to-close are constructed by taking the differences of the log-prices available for each commodity at 14.00. Taking the 'official' end-of-day closing price for each commodity would have distorted the returns comovement measurement because markets do not all close concurrently and have different rules for determining their 'settlement' price.

As shown in Table 1, Soybean meal, Rough Rice, Oats in the 'Grains' sector; Lumber and Orange Juice in the 'Softs' sector; and Pork Bellies in the 'Livestock' sector are commodities not included in any of the indices. These commodities are always included in the 'off-index'. Note that other commodities do appear in some indices but not in others. These are RBOB Unlead Gas (not in Thomson Reuters), Soybean oil (not in S&P GSCI), Cocoa (not in Dow Jones UBS), feeder cattle (only in S&P GSCI) and Platinum (only in Thomson-Reuters). As described later, these commodities are included in the corresponding 'off-index' as well.

2.2 Univariate regressions

We begin by testing the commodity comovement hypothesis through a univariate regression, described by Prediction 1. If Prediction 1 holds, index commodities display an increase in comovement. Likewise, off-index commodities are not expected to show increased comovement with commodities in the index. This is not to say that off-index commodities should show no comovement with the index. Instead, comovement between off-index commodities and the index should be less than comovement between the index commodities and the index, and comovement between off-index commodities and the index is not expected to increase post-2005.

To test this, for each commodity we estimate the univariate regression

$$R_{i,t} = \alpha_i + \beta_i R_{CMI,t} + \nu_{i,t}$$

separately for the period before 1 January 2005 and after; where $R_{i,t}$ is the return of commodity i at time t and $R_{CMI,t}$ is the return of S&P GSCI or DJ-UBS CI or the Thomson-Reuters CI. We estimate the change in the slope coefficient before and after January, 2005 ($\Delta\beta = \beta_{post} - \beta_{pre}$) and the change in the R^2 ($\Delta R^2 = R^2_{post} - R^2_{pre}$) using three data frequencies (daily, weekly, and monthly). The first period goes from April 9, 1998 to December 30, 2004, or 1,659 trading days. The second period goes from January 3 2005 to March 24, 2011, or 1,564 days. Note that the chosen event date roughly splits the sample in to series of the same length. The return of the index is calculated using the index specific weights as in Table 1 and all commodities except the one being tested. Note that the sum of the weights does not equal 1. This affects the value of the β coefficient but does not affect the test statistics on $\Delta\beta$.[7]

[6] Data are at 1-minute frequency covering the trading hours of the CME and CBoT. Globex trading information (i.e. outside the markets trading hours) is at our disposal for major commodities only.

[7] Rescaling all weights so that they sum up to 1 while keeping their proportions to the remaining weights obviously leads to a different β coefficient; the test value under the null that $\Delta\beta$ is zero is unaffected.

	Commodity	Ticker	Exchange[a]	S&P GSCI	DJ-UBS	Th.R
Energy	WTI crude oil	CL	NYMEX	40.6	15	5.88
	Heating Oil	HO	NYMEX	5.3	4.5	5.88
	RBOB unlead gas	XRB	NYMEX	4.5	4.1	-
	Natural Gas	NG	NYMEX	7.6	16	5.88
Grains	Corn	C	CME Group	3.6	6.9	5.88
	Soybeans	S	CME Group	0.9	7.4	5.88
	Chicago wheat	W	CME Group	3.7	3.4	5.88
	Soybean oil	BO	CME Group	-	2.9	5.88
	Soybean meal	SM	CME Group	-	-	-
	Rough rice	RR	CME Group	-	-	-
	Oats	O	CME Group	-	-	-
Softs	Coffee	KC	ICE	0.5	2.7	5.88
	Cotton	CT	ICE	0.7	2.2	5.88
	Sugar	SB	ICE	2.1	2.8	5.88
	Cocoa	CC	ICE	0.2	-	5.88
	Lumber	LB	CME Group	-	-	-
	Orange Juice	JO	ICE	-	-	-
Livestock	Feeder cattle	FC	CME Group	0.3	-	-
	Lean hogs	LH	CME Group	0.8	2.5	5.88
	Live cattle	LC	CME Group	1.6	4.1	5.88
	Pork bellies	PB	CME Group	-	-	-
Metals	Gold	GC	NYMEX	1.5	6.1	5.88
	Silver	SI	NYMEX	0.2	2.4	5.88
	Copper	HG	NYMEX	2.6	6.7	5.88
	Platinum	PL	NYMEX	-	-	5.88

Table 1: Commodity futures traded in the United States (sector, ticker, exchange, and index weight data). The S&P GSCI and DJ-UBS CI also include commodities traded in London, which were not part of our analysis. The indices weights are taken as of 2008. The Thomson Reuters is an equally weighted index.

[a] NYMEX denotes the New York Mercantile Exchange; CME Group the Chicago Mercantile Exchange.

This differs both from Barberis et al. (2005) and previous analyses of the commodity financialization question (e.g. Tang and Xiong (2012)) in that we test Prediction 1 for each commodity. Barberis et al. (2005) test the *mean* of βs of the univariate regressions of all stock included in the index before and after inclusion. Tang and Xiong (2012) rely on a pooled regression across commodities; the effect of financialization is captured by introducing two additional coefficients measuring the trend after 2004 for all commodities and for index commodities. Given the small sample of 25 commodities, it is possible to analyse and discuss whether Prediction 1 holds case by case, i.e. for every commodity, as is done here.

Table 2 reports the change in the slope coefficient and the change in R^2 of each of the commodity indices for the 25 commodities studied (using daily returns). A number of interesting features are evident. First, we observe a statistically significant increase in comovement between index non-energy commodities and heavy-energy indices[8] (positive $\Delta\beta$). We also observe a statistically significant increase in comovement between index non-energy commodities and the equally weighted index (Thomson reuters CI) although it is less pronounced. Moreover, the increase in comovement between off-index non-energy commodities and the index is smaller (soybean meal, rice, oats) or statistically less significant (lumber and orange juice and pork bellies) than the index commodities. These findings are in line with that expected from Prediction 1.

However, in contrast to the index non-energy commodities, there is no increase in comovement between the index energy commodities and the index (the $\Delta\beta$'s observed are generally negative and not statistically significant). This is potentially an indication of the presence of financialization of specifically energy commodities pre-2005. This could also be an artefact of the fact that for the S&P GSCI and UBS Dow Jones CI, energy commodities represent by far the most weight in the index (58% and 40%, respectively); we are essentially regressing the energy commodities on themselves. Thus the regression coefficient of one energy commodity will be severely conditioned to returns on the remaining three commodities. For these two energy-heavy indices we should thus not expect significant changes in comovement. This is not the case for the equally weighted commodity index and begs an alternative explanation. Energy commodities display also high correlations since they are either related to the same product (oil) or to the same market-wide shock (e.g. economic factors, wars, geopolitical patterns). The observed results are thus in line with what we would expect. For the energy-heavy indices, we do not expect to see big changes in their correlation structure due to an increase in index trading. For the third index, we do not expect to see big changes in their dependence structure due to their already significantly high correlation level.

Focusing back on the non-energy commodities, our results show, consistent with the finding in Tang and Xiong (2012) who uses Oil return as a proxy for index return, that index non-energy commodities increased their comovement with the commodities in the index (oil being the most important

[8]Energy commodities have by far the greatest weight in S&P GSCI and DJ-UBS CI (58% and 40%, respectively - Table 1), thus making them 'heavy-energy' indices. The Thomson-Reuters CI is an equally-weighted index and thus does not assign preferential weight to the energy sector.

in the index) as an effect of financialization. Having estimated a coefficient for each commodity, we cannot comment on the magnitude of the increase for global index and global off-index commodities.

To confirm this and also isolate the comovement effect produced by Heating Oil, RBOB Unlead Gas and Natural Gas, we repeate the analyses keeping Oil as the only energy commodity. When considering exclusively oil for the energy commodities (Table 3), we obtain results in line with what we find in the literature (Nazlioglu et al. (2013) and Sadorsky (2014)). Namely, we observe a statistically significant increase in comovement between non-energy commodities and "oil", both index and off-index. However, for off-index commodities, the increase in comovement is smaller. Specifically, for the S&P GSCI we see a general beta increase across commodities. Note also that now Oil displays a significant increase in comovement with respect to the other index non-energy commodities ($\Delta\beta = 3.40$). For the other two indices, which have a lower or equal allocation in Oil, we see that except for Lumber and Oats (in one case only for both), all off-index commodities do not display significant increase in β or, for Pork Bellies, there is a significant decrease. Note also that Feeder Cattle is off-index for these two cases and does not display any slope change.

Tables 4 and 5 report the results for weekly and monthly regressions, respectively.[9] As the data frequency decreases (from daily to weekly to monthly), the financialization effect is observed to weaken ($\Delta\beta$'s decrease and become less statistically significant). This effect is reported in Barberis et al. (2005) as well. The predictions stemming from the alternative (sentiment-based) theories of comovement imply a weakening of the β increase effect at sufficiently low frequencies. This is because eventually index trader sentiment will revert or any slowly-diffusing information will be incorporated and in the end returns will be more closely tied to fundamentals.

To further supplement this analysis, it would have been interesting to test how commodity index recompositions affect commodity prices and comovement. In a way, our use of three different indices, each with different commodity weightings, already explores this question, as each index hosts a unique composition. However we are not able to directly investigate the impact of time-varying, continuous recomposition. To partially address this question, we performed a series of tests in which the weightings of the commodities in the S&P GSCI were slightly changed each year (while keeping the total weight constant). The results obtained with this index recomposition were not materially different from the original results (data available upon request); the impact of index recomposition is thus likely to be minimal.

In summary, starting in 2005, comovement between index non-energy commodities and the index has significantly increased, which suggests that index commodities have become less segmented (indicating financialization). This result is in line with the findings presented in Tang and Xiong (2012). For off-index commodities, either no change or a significant decrease in comovement is observed. Thus after 2005, two distinct groups of commodities emerged - financialized commodities (those included in the major indices) and non-financialized commodities (off-index commodities). This is comparable to the finding of two distinct groups of stocks by Barberis et al. (2005) - stocks included in

[9]Weekly and monthly series are generated by taking the price in the last minute of the week and month, respectively.

the S&P 500 and those not on the index. While Barberis and colleagues' analysis was concerned with the impact of being included in the S&P index (which was found to increase return correlations with the index returns), here the focus was examining the difference in correlations between pre-2005 and post-2005 returns. Index commodities exhibit an increase in comovement with the index post-2005 while non-index commodities do not, providing additional evidence for the financialization impact of index investing on the commodities market.

Table 2: Univariate regressions

	S&P GSCI		DJ-UBS CI		Th. Reuters CI	
Commodity	$\Delta\beta$	ΔR^2	$\Delta\beta$	ΔR^2	$\Delta\beta$	ΔR^2
WTI crude oil	-0.27	0.10	0.07	0.20	0.12	0.28
Heating Oil	-0.25	0.01	-0.35	0.11	-0.13	0.24
RBOB un. gas	-0.06	0.01	-0.12**	0.11	-0.21	0.18
Natural Gas	-0.13	0.00	-0.64	-0.01	-0.45	0.02
Corn	0.52	0.15	0.54	0.19	0.39	0.21
Soybeans	0.55	0.21	0.57	0.26	0.37	0.29
Chicago wheat	0.46	0.11	0.46	0.14	0.36	0.16
Soybean oil	0.67	0.31	0.65	0.37	0.53	0.39
Soybean meal	0.42	0.12	0.36	0.17	0.12*	0.17
Rough rice	0.16	0.03	0.16	0.04	0.07	0.05
Oats	0.37	0.09	0.30	0.11	0.02	0.11
Coffee	0.36	0.08	0.43	0.11	0.29	0.16
Cotton	0.37	0.10	0.39	0.12	0.43	0.17
Sugar	0.49	0.10	0.55	0.13	0.52	0.16
Cocoa	0.32	0.08	0.35	0.1	0.31	0.14
Lumber	0.12**	0.01**	0.12**	0.01*	0.11	0.02**
Orange Juice	0.13*	0.01**	0.12**	0.02**	0.09	0.02*
Feeder cattle	0.10	0.03	0.06**	0.02	0.01	0.02
Lean hogs	0.07	0.01	0.01	0.01	-0.10	0.01
Live cattle	0.14	0.05	0.12	0.05	0.12	0.07
Pork bellies	0.04	0.00	-0.03	0.00	-0.28	-0.01
Gold	0.26	0.13	0.27	0.14	0.22	0.19
Silver	0.59	0.19	0.63	0.21	0.63	0.24
Copper	0.73	0.25	0.80	0.25	0.86	0.31
Platinum	0.38	0.16	0.39	0.20	0.40	0.26

Table 3: Univariate regressions - Oil only

	S&P GSCI		DJ-UBS CI		Th. Reuters CI	
Commodity	$\Delta\beta$	ΔR^2	$\Delta\beta$	ΔR^2	$\Delta\beta$	ΔR^2
WTI crude oil	3.40	0.25	1.59	0.29	1.03	0.30
Corn	0.68	0.16	0.61	0.22	0.29	0.19
Soybeans	0.70	0.22	0.62	0.30	0.22	0.26
Chicago wheat	0.61	0.12	0.44	0.16	0.30	0.16
Soybean oil	0.84	0.32	0.66	0.41	0.38	0.35
Soybean meal	0.54	0.13	0.10	0.17	-0.07	0.15
Rough rice	0.20	0.03	0.09	0.05	0.00	0.05
Oats	0.46	0.09	0.00	0.11	-0.19	0.10
Coffee	0.44	0.08	0.48	0.13	0.15	0.15
Cotton	0.50	0.11	0.59	0.17	0.49	0.19
Sugar	0.63	0.10	0.79	0.15	0.55	0.15
Cocoa	0.42	0.08	0.45	0.13	0.29	0.14
Lumber	0.18**	0.01 **	0.22	0.02	0.14	0.02
Orange Juice	0.17**	0.01 **	0.19	0.02	0.11	0.02
Feeder cattle	0.13	0.03	0.08	0.02	0.00	0.01
Lean hogs	0.10	0.01	-0.12	0.01	-0.17	0.01
Live cattle	0.19	0.05	0.20	0.07	0.15	0.07
Pork bellies	0.06	0.00	-0.22	0.00	-0.45	-0.02
Gold	0.34	0.14	0.33	0.17	0.21	0.19
Silver	0.76	0.20	0.83	0.25	0.66	0.24
Copper	0.94	0.26	1.27	0.31	0.94	0.31
Platinum	0.50	0.18	0.52	0.26	0.42	0.27

Reported are the β and R^2 values pre and post January 1, 2005 and their difference. Positive significant differences from zero at the 10% and 5% levels in two-sided tests, respectively are denoted by * and **. Light gray cells denote that the difference is not significant at 10% level or higher; dark gray cells denote negative significant differences at 10% level. Rest significant at 1% level. Standard errors are computed using the White robust covariance matrix. Commodities in bold are those not included in any index.

Table 4: Univariate regressions - weekly returns

	S&P GSCI		DJ-UBS CI		Th. Reuters CI	
Commodity	$\Delta\beta$	ΔR^2	$\Delta\beta$	ΔR^2	$\Delta\beta$	ΔR^2
WTI crude oil	-0.31	0.09	0.03	0.21	0.04	0.27
Heating Oil	-0.24	0.02	-0.31	0.17	-0.07	0.30
RBOB unlead gas	0.02	0.07	-0.06	0.15	-0.20	0.20
Natural Gas	-0.02	0.03	-0.56	0.02	-0.24	0.06
Corn	0.42	0.10	0.44	0.14	0.26	0.15
Soybeans	0.58	0.24	0.55	0.27	0.33	0.32
Chicago wheat	0.18	0.02	0.18	0.04	-0.02	0.04
Soybean oil	0.77	0.33	0.76	0.42	0.70	0.48
Soybean meal	0.37	0.13	0.24*	0.16	-0.07	0.15
Rough rice	0.27	0.03	0.23	0.05*	0.02	0.06*
Oats	0.41	0.10	0.41	0.14	0.21	0.15
Coffee	0.57	0.13	0.60	0.17	0.40**	0.23
Cotton	0.40	0.10	0.47	0.14	0.55	0.21
Sugar	0.20	0.03	0.24	0.06*	0.01	0.07
Cocoa	0.50	0.10	0.62	0.15	0.55	0.21
Lumber	0.10	0.01	0.17	0.02	0.23	0.04*
Orange Juice	0.21	0.02	0.23*	0.04	0.14	0.04
Feeder cattle	0.18	0.05*	0.16	0.05*	0.17*	0.05*
Lean hogs	0.07	0.00	0.03	0.00	0.12	0.00
Live cattle	0.17	0.07	0.17*	0.08	0.22	0.10
Pork bellies	0.11	0.00	0.00	0.00	-0.34	-0.02
Gold	0.43	0.23	0.43	0.24	0.36	0.28
Silver	0.78	0.27	0.81	0.30	0.73	0.32
Copper	0.71	0.27	0.76	0.28	0.66	0.31
Platinum	0.49	0.27	0.52	0.33	0.44	0.36

Table 5: Univariate regressions - monthly returns

	S&P GSCI		DJ-UBS CI		Th. Reuters CI	
Commodity	$\Delta\beta$	ΔR^2	$\Delta\beta$	ΔR^2	$\Delta\beta$	ΔR^2
WTI crude oil	0.03	0.04	0.15	0.16	-0.19	0.15
Heating Oil	-0.13	-0.01	-0.12	0.15	-0.09	0.26**
RBOB unlead gas	0.00	-0.01	0.19	0.19	-0.28	0.17
Natural Gas	-0.31	-0.01	-1.04	-0.02	-0.95	0.00
Corn	0.56	0.21*	0.68	0.29	0.53*	0.29
Soybeans	0.70	0.25	0.69	0.32	0.57*	0.39
Chicago wheat	0.01	-0.03	0.00	-0.05	0.05	0.01
Soybean oil	0.71	0.29	0.77	0.45	1.00	0.55
Soybean meal	0.55	0.18	0.35	0.23*	0.19	0.27
Rough rice	0.41	-0.01	0.39	0.03	0.15	0.06
Oats	0.54	0.19	0.66	0.31	0.75	0.34
Coffee	0.44*	0.07	0.59	0.13	0.58*	0.19**
Cotton	0.33	0.08	0.38	0.12	0.37	0.18
Sugar	0.23	0.03	0.37	0.05	0.46	0.07
Cocoa	0.15	0.03	0.37	0.06	0.52	0.14
Lumber	0.40	0.13	0.60	0.13*	0.67*	0.18*
Orange Juice	0.46	0.07	0.66	0.16	0.86	0.14
Feeder cattle	0.04	0.01	0.03	0.01	0.04	0.02
Lean hogs	-0.27	0.04	-0.28	0.03	0.04	0.01
Live cattle	-0.06	-0.02	-0.04	-0.01	-0.10	-0.01
Pork bellies	-0.38	0.04	-0.57	0.03	-0.91	-0.01
Gold	0.20	0.08	0.19	0.08	0.18	0.12
Silver	0.60	0.19	0.78	0.26	0.82	0.29
Copper	0.67	0.24*	0.68	0.20*	0.67	0.22
Platinum	0.26	0.14	0.33	0.22*	0.26	0.28

Reported are the β and R^2 values pre and post January 1, 2005 and their difference. Positive significant differences from zero at the 10% and 5% levels in two-sided tests, respectively are denoted by * and **. Light grey cells denote that the difference is not significant at 10% level or higher; dark grey cells denote negative significant differences at 10% level. Rest significant at 1% level. Standard errors are computed using the White robust covariance matrix. Commodities in bold are those not included in any index.

2.3 Bivariate regressions

We continue our empirical analysis by testing the commodity comovement hypothesis using a bivariate regression, described by Prediction 2. As discussed in Section 2, the βs calculated from the univariate regressions are not clean measures of sentiment shock. Some of the covariation will be caused by news related to the commodity fundamentals (the traditional explanation for comovement) in addition to the covariation caused by commodities being in the same category or investment habitat (the alternative sentiment-based theories for comovement tested for here). In the bivariate regression, two independent variables are used: one for the returns of index commodities ($\beta_{i,IND}$) and one for the returns of off-index commodities ($\beta_{i,OFF}$). The inclusion of an 'off-index' commodity index acts to control for such news about commodity fundamentals, thus making $\beta_{i,IND}$ a much cleaner measure of sentiment-based comovement. The variable representative of the off-index commodity returns was created by equally weighting the returns of all the off-index commodities studied here. As a reminder, the six commodities not included in the indices are always included in the 'off-index': Soybean meal, Rough Rice, Oats in the 'Grains' sector; Lumber and Orange Juice in the 'Softs' sector; and Pork Bellies in the 'Livestock' sector. The commodities that are included in some indices but not in others[10] were included in the 'off-index' corresponding index.

If Prediction 2 holds, returns of index commodities will tend to display an increase in comovement with the other index commodities and a decrease in comovement with the off-index, when controlling for the returns of off-index commodities. In other words, the presence of commodity i in the index (IND) increases the covariance of its return with the return on the index and hence also its beta loading on the index return, $\beta_{i,IND}$. At the same time, it decreases the covariance of its return with the return on the off-index commodity index (OFF) and hence also its beta loading on the off-index return, $\beta_{i,OFF}$.

To test this, for each commodity we estimate the bivariate regression:

$$R_{i,t} = \alpha_i + \beta_{i,IND} R_{IND,t} + \beta_{i,OFF} R_{OFF,t} + \nu_{i,t}$$

where

$$R_{OFF,t} = \frac{1}{m} \sum_{k=1}^{m} R_{k,t}, \quad k \notin IND$$

and IND denotes the set of commodities included in an index. For ease of exposition, we restricted our analysis to a daily return frequency.

The results of this bivariate regression are presented in Table 6. The first feature of note is that there is a statistically significant increase in comovement between index non-energy commodities and the index (β_{IND} is generally increasing), as seen in the univariate regressions. Examining the off-index regression coefficients for the same commodities, we see that while β_{OFF} is significant and increasing for the indices with heavy Oil allocation (S&P GSCI and UBS-DJ CI), no significant

[10] RBOB Unlead Gas (not in Thomson Reuters), Soybean oil (not in S&P GSCI), Cocoa (not in Dow Jones UBS CI), feeder cattle (only in S&P GSCI) and Platinum (only in Thomson-Reuters CI)

change in β_{OFF} is observed for the equally weighted Thomson-Reuters index. This may be because the indices heavily weighted in oil cannot adequately capture price changes in non-energy commodities (given that 40-60% of the index is comprised of energy commodities); these price changes are instead captured by the off-index returns. Thus when comovement between the index commodity and the index increases, the comovement with off-index commodities also increases. The increasing β_{OFF} observed for the S&P GSCI and DJ-UBS CI indices is therefore most likely an artefact of the composition of the index and not an actual increase in correlation with the off-index returns. When an equally weighted index is used, only the comovement between an index commodity and the index increases; the comovement with the off-index commodities does not change.

The second major finding is that, in contrast to the results of the univariate regression, commodities that are off-index do not display an increase in β with the index and also do not show a significant increase in comovement with the off-index returns. Hence, off-index commodities do not display any change in comovement. This suggests that off-index commodities remained partially segmented from each other post-2005 while the index commodities experienced a significant increase in comovement. This provides evidence for the financialization of U.S. commodity markets post-2005.

Finally, as observed in the univariate regressions, the index energy commodities do not exhibit the same behaviour as the index non-energy commodities. While for the index non-energy commodities generally exhibit significant, positive $\Delta\beta_{IND}$ values, the $\Delta\beta$'s for index energy commodities are generally negative and not statistically significant. This is potentially an indication of the presence of financialization of specifically energy commodities pre-2005. However, this observation could also be an artefact of the fact that for the S&P GSCI and DJ-UBS CI, energy commodities represent by far the most weight in the index (58% and 40%, respectively). This heavy energy weighting could significantly skew any measurement of covariation between an energy commodity and the index. Thus the apparent differences between energy and non-energy commodities must be interpreted with care.

In summary, the bivariate regressions demonstrate that for those non-energy commodities that are included in a commodity index, comovement with the index increased after 2005. Off-index commodities, on the other hand, do not exhibit increased comovement with the index or the 'off-index' commodities index. These findings are in line with that expected from Prediction 2. However, according to Prediction 2, for index commodities we should also observe a decreasing comovement with the 'off-index' (decreasing β_{OFF}). We observe either an increase in β_{OFF} (for the oil-heavy commodities, explained above) or no change. Thus while some of these results hint at the influence of financialization, we cannot confidently interpret the bivariate off-index findings as definitive evidence for or against financialization.

Table 6: Bivariate regressions

	S&P GSCI		DJ-UBS CI		Th. Reuters CI	
Commodity	$\Delta\beta'^{ind}$	$\Delta\beta'^{off}$	$\Delta\beta'^{ind}$	$\Delta\beta'^{off}$	$\Delta\beta'^{ind}$	$\Delta\beta'^{off}$
WTI crude oil	-0.44	0.40	-0.07	0.46	0.24	-0.21
Heating Oil	-0.24	0.07	-0.36	0.22	-0.06	-0.12
RBOB un. gas	-0.05	0.12	-0.12	0.22	-0.32	0.29
Natural Gas	-0.12	-0.17	-0.53	-0.14	-0.30	-0.23
Corn	0.18	0.22	0.29	0.20	0.33	0.09
Soybeans	0.20	0.11	0.31	0.16	0.32	0.07
Chicago wheat	0.16	0.26	0.25	0.22	0.33	0.05
Soybean oil	0.50	0.10	0.51	0.04	0.47	0.10
Soybean meal	0.21	0.07	0.27	0.04	0.13	-0.02
Rough rice	0.05	0.02	0.09	0.01	0.11	-0.05
Oats	0.15	0.08	0.19	0.05	-0.05	0.14
Coffee	0.21	0.23	0.29	0.16	0.22	0.13
Cotton	0.17	0.31	0.21	0.34	0.38	0.09
Sugar	0.36	0.24	0.43	0.22	0.48	0.06
Cocoa	0.21	0.13	0.31	0.03	0.43	-0.21
Lumber	0.04	0.13	0.05	0.08	0.07	0.06
Orange Juice	0.06	0.13	0.08	0.07	0.05	0.07
Feeder cattle	0.10	-0.02	0.06	-0.01	-0.01	0.03
Lean hogs	0.02	-0.39	0.02	-0.49	0.23	-0.60
Live cattle	0.08	0.09	0.04	0.09	0.10	0.04
Pork bellies	-0.01	-0.02	-0.07	-0.02	-0.32	0.06
Gold	0.17	0.14	0.19	0.14	0.28	-0.11
Silver	0.37	0.41	0.42	0.42	0.64	-0.02
Copper	0.54	0.38	0.61	0.40	0.80	0.11
Platinum	0.28	0.19	0.30	0.19	0.38	0.03

Bivariate regressions. Positive significant difference from zero at 10% and 5% level in two-sided test are denoted by $*$ and $**$ respectively. Dark grey cell negative and significant at the 10%; light grey cell denote difference is not significant at the 10%. Rest significant at 1% level. Commodities in bold are those not included in any index.

2.4 Alternative explanations and robustness checks

We now consider some alternative explanations for the results of the univariate and bivariate regressions. In their analysis of sentiment-based comovement and the S&P 500, Barberis et al. (2005) describe two possible alternative explanations for the observed change in comovement after a stock is included in the S&P index: non-trading effects and characteristic and demand effects. Here we discuss how these alternative explanations could impact our analysis of the commodities market.

Non-trading effects

The first alternative explanation we explore is the potential impact of 'non-trading effects'. The basic premise of this argument is that since index commodities are generally more liquid and more frequently traded than off-index commodities, the correlation results might have some spurious upward bias. To see this, assume that some important commodity market information is revealed toward the end of the trading day. It might be the case that index commodities futures, being more liquid, will trade again before the end of the day and thus assimilate such information. On the contrary, off-index commodities, trading less often, might not trade again before the end of the day. As a result, returns $R_{i,t}$ on off-index commodities will not reflect such news and the regression of $R_{i,t}$ on $R_{IND,t}$ would produce an artificially low slope coefficient.

We address this possible issue in two ways: by restricting the trading time analysed and investigating changes in trading activity. First, recall that in our analyses we considered the day trading time going from 10:30 to 14:00. This is the overlapping period in which all commodities futures trade across different US markets. Restricting ourselves to the overlapping trading day should reduce the occurrence of non-trading since it is likely that off-index commodities also trade frequently during this time. Had we chosen a different period, non-trading effects would be surely observed for commodities with a narrower trading day. In addition, in recent years, trading across the clock has become possible and has consistently increased in volume. This means that in our setup, any news arriving after 14:00 will be reflected in next day's price thus further increasing the likelihood that all commodities trades on that information.

To explicitly test for the impact of 'non-trading' effects, we investigate the average daily contract volume of each commodity pre- and post-2005. According to the non-trading effect hypothesis, we observe an increase in β simply because an index commodity trades more frequently than the off-index one. In order to test this hypotheses, we consider, for each commodity, the average daily contract volume from 1998 to 2011 as a measure of how frequently the commodity is traded. Table 7 reports the direction of the change in average daily contract volume post 2005. Consider the case of Grains. All of the commodities in this sector but Oats show an increase in the daily contract volume in 2005-2011 compared to 1998-2004. Although all of the grain commodities show an increase in daily contract volume, the index commodities (corn, soybeans, wheat, and soybean oil) experienced an increase in comovement after 2005 while the off-index commodities (soybean meal, rough rice and oats) did not experience an increase in comovement with the index. We consider this to be evidence of commodity (grain) financialization; non-trading effects cannot explain the results of our daily-frequency regressions.

Grains	Corn	$\nearrow$	Energy	WTI crude oil	$\nearrow$
	Soybeans	$\nearrow$		Heating Oil	$\nearrow$
	Chicago wheat	$\nearrow$		RBOB unleaded gas^{a}	
	Soybean oil	$\nearrow$		Natural Gas	$\nearrow$
	Soybean meal	$\nearrow$			
	Rough rice	$\nearrow$	Livestock	Feeder cattle	$\nearrow$
	Oats	$\searrow$		Lean hogs	$\nearrow$
				Live cattle	$\nearrow$
Softs	Coffee	$\nearrow$		**Pork bellies**	$\searrow$
	Cotton	$\nearrow$			
	Sugar	$\nearrow$	Metals	Gold	$\nearrow$
	Cocoa	$\nearrow$		Silver	$\nearrow$
	Lumber	$\nearrow$		Copper	$\nearrow$
	Orange Juice	$\nearrow$		Platinum	$\nearrow$

Table 7: Change in daily contract volume after 2005. $\searrow$ indicates a decrease in average daily contract volume in 2005-2011 compared to 1998-2004. $\nearrow$ indicates an increase in average daily contract volume. Commodities in bold are those not included in any index.

a RBOB unleaded gas daily contract volume series is incomplete.

Source: Quandl.

Characteristic and demand effects

When considering the case of stock inclusion in the S&P 500 index, characteristic and demand effects describe the idea that the increase in comovement could arise if the stocks in the S&P 500 index differ from other stocks in terms of some characteristic, and the stocks Standard & Poor's chooses to include are those increasingly demonstrating that characteristic. If that characteristic happens to be associated with a cash-flow factor, the presence of the common characteristic could explain the increase in comovement of the returns. For stocks, one such characteristic is size (as stocks Standard & Poor's chooses tend to be growing in size prior to inclusion). In our case, one could argue that index commodities are those commodities that better reflect the entire economy and cover the entirety of the world production.

Barberis et al. (2005) address this competing explanation with a matching exercise. For each event stock included in the S&P during the sample period, they search for a "matching" stock, drawn from the same industry as the event stock and in the same size range at the time of inclusion, but which is not included in the index. We perform a similar 'matching' exercise by comparing the returns of soybeans (included in the main indices), soybean oil (again, indexed), and soybean meal (off-index). Soybeans and their derivative products are chosen as a case study given that the processing of soybeans necessarily results in the production of both soybean meal and soybean oil. Soybean meal is a solid residue by-product, flour, created after grinding the soybean to extract soybean oil. Thus the total supply of each derivative commodity will always remain constant relative to the other, meaning that prices of both soybean meal and oil are primarily influenced by the total supply/demand for soybeans.[11] This should make the three soybean products good matching commodities - they likely exhibit the same financial characteristics.

Recalling the results in Table 2, we observe that the off-index commodity, soybean meal, exhibits much smaller shifts in betas than the indexed soybeans and soybean oil. Table 3 reports results that are statistically more relevant. Soybeans and soybean oil exhibit increase in comovement whereas soybean meal beta increase is statistically insignificant (for DJ-UBS CI and Thomson Reuters CI). Given that these three commodities can be considered matching, this finding suggests that the increase in comovement post-2005 is not entirely due to characteristic effects. We thus consider this example to be possible evidence of commodity (soybean) financialization.

Robustness checks

To conclude this Section, we test whether our methodology delivers results consistent with the sentiment- or friction-based views also in a period pre-financialization. We consider the interval from 1998 to 2004 (inclusive), split the time period into two equally long time series and re-run the regressions to test Prediction 1 as in Section 2.2. This time period is that prior to the boom and bust of the commodity futures prices, i.e. before the effects of financialization were visible. We therefore expect no changes in comovement.

Results from this additional test are displayed in Table 8. For nearly all commodities, no significant change in beta nor in R^2 is observed between the two periods. This is in contrast to the findings for the pre- and post- 2005 periods, where a significant increase in beta is generally observed for the index commodities. The change in comovement interpreted as evidence of financialization is thus not present in this prior time period. These results further corroborate our previous findings: changes in intra-commodity comovements cannot be justified as driven by changes in fundamental values.

3 Realized betas

To conclude the empirical analysis, we present a new approach to validate Prediction 1. Instead of regressing Equation 1 on daily returns, we use price information at intra-day frequency and estimate

[11] A move in the price of one derivative commodity against the other is indicative of high demand for one of the products.

	S&P GSCI		DJ-UBS CI		Th. Reuters CI	
Commodity	$\Delta\beta$	ΔR^2	$\Delta\beta$	ΔR^2	$\Delta\beta$	ΔR^2
WTI crude oil	-0.06	0.14	0.22	0.15	0.14	0.09
Heating Oil	0.06	0.03	0.02	0.08	0.10	0.06
RBOB unlead gas	0.16	-0.02	0.15	0.05	0.09	0.03
Natural Gas	0.58	0.12	0.57	0.08	0.25	0.03
Corn	0.03	0.00	-0.04	-0.01	-0.18	-0.04
Soybeans	0.04	0.00	0.00	-0.01	-0.04	-0.05
Chicago wheat	0.06	0.01	-0.05	-0.02	-0.18	-0.05
Soybean oil	0.08	0.00	0.09	0.01	0.12	0.02
Soybean meal	0.01	0.00	-0.01	-0.02	-0.08	-0.05
Rough rice	0.02	0.00	-0.02	0.00	0.01	0.00
Oats	0.00	0.00	-0.08	-0.02	-0.14	-0.05
Coffee	0.18	0.00	0.21	0.01	0.33	0.03
Cotton	0.12	0.01	0.10	0.01	0.17	0.01
Sugar	-0.01	0.00	-0.07	0.00	-0.12	-0.01
Cocoa	0.07	0.00	0.10	0.01	0.02	0.00
Lumber	-0.08	0.00	-0.12	0.00	-0.13	-0.01
Orange Juice	0.04	0.00	0.03	0.00	0.06	0.01
Feeder cattle	-0.03	0.00	0.00	0.00	0.04	0.00
Lean hogs	0.01	0.00	0.00	0.00	0.04	0.01
Live cattle	-0.03	0.00	0.00	0.00	0.02	0.00
Pork bellies	-0.08	0.00	-0.13	0.00	-0.30	-0.01
Gold	0.07	0.02	0.10	0.04	0.12	0.05
Silver	0.10	0.01	0.17	0.03	0.30	0.05
Copper	0.09	0.01	0.18	0.03	0.28	0.05
Platinum	0.04	0.01	0.07	0.02	0.02	0.03

Table 8: Robustness checks: univariate regressions run during the 'prefinancialization' 1998-2004 period as a control. The 1998-2004 period was split into two equally long time series to calculate the change in β and R^2. Light gray cells denote that the difference is not significant at 10% level or higher; dark gray cells denote negative significant differences at 10% level. Commodities in bold are those not included in any index.

directly the so called ‘realized beta’. This is possible because of the availability of price information at intra-day frequency.

The concept of realized betas was first introduced by Bollerslev and Zhang (2003). They carried out a large-scale estimation of the Fama-French three-factor model using 5-minute data on 6,400 stocks over a period of 7 years. They show that using high-frequency data can improve the pricing accuracy of asset pricing models. In a related paper Andersen et al. (2006) investigates the time variation in realized variances, covariances and betas using daily returns to construct quarterly realized measures. They found evidence of strong persistence in the variance and covariance process, but less persistency in the beta process. This indicates that realized volatility and realized covariance are fractionally cointegrated.[12]

We calculate the realised betas as follows: defining R_{i,t_k} and R_{CMI,t_k} the intra-day returns at time $t_k = 1, \dots, m$, of the i-th commodity and the commodity index, respectively, and with m the number of intra-day returns, the realized beta reads:

$$\beta_{i,t} = \frac{\sum_{k=1}^{m} R_{i,t_k} R_{CMI,t_k}}{\sum_{k=1}^{m} R_{CMI,t_k}^2}. \tag{5}$$

The numerator in Equation (5) is the classical estimator of the realized covariance, the denominator is the classical estimator of the realized variance.

Before discussing the results, an important caveat to this analysis is that an increase in realized betas indicates an increase in return comovement in the absence of micro-structure noise. In the presence of jumps and micro-structure noise (e.g. stale prices, asynchronous trading) the realised variance estimator (denominator of the realised beta) is not longer efficient. Additionally, as the sample frequency increases, the realized covariance estimator is biased toward zero, the so-called Epps (1979) effect. Following Hansen et al. (2014), in this paper we rely on the multivariate kernel estimator proposed in Barndorff-Nielsen et al. (2011). This consistent estimator is guaranteed to be positive semi-definite and is robust to measurement noise and to non-synchronous trading.

With this proviso, results are presented in Tables 9 and 10 using all commodities and only Oil as energy commodity, respectively. We observe a statistically significant increase in realized betas of index non-energy commodities and index, and a smaller or non significant increase in realized betas of off-index non-energy commodities and index. These results support the regressions on daily returns and indicate that on average, realized betas of index commodities increase after 2004. Moreover, this increase in realised betas tells us that the increase in return comovement can be seen not only at a daily frequency, but even at intra-day levels. This provides further evidence that the increase in comovement observed is not related to changes in fundamentals and instead represents additional evidence of financialization as the driver of the observed increase in commodity return comovement.

[12]Other related studies include Barndorff-Nielsen and Shephard (2004), who derived asymptotic results for realized beta and Dovonon et al. (2013), who established a theory for bootstrapping inference.

Table 9: Average Realized Betas

	S&P GSCI	DJ-UBS CI	Th.Reuters CI
Commodity	$\Delta\beta$	$\Delta\beta$	$\Delta\beta$
WTI crude oil	-0.14	0.16	0.28
Heating Oil	-0.02	-0.11	0.17
RBOB unlead gas	0.12	0.02	0.07**
Natural Gas	0.35	0.21	0.27
Corn	0.33	0.39	0.49
Soybeans	0.28	0.31	0.25
Chicago wheat	0.21	0.21	0.23
Soybean oil	0.25	0.22	0.19
Soybean meal	0.20	0.14	0.06**
Rough rice	-0.01	-0.01	-0.03
Oats	0.01	-0.10	-0.24
Coffee	0.16	0.20	0.29
Cotton	0.16	0.20	0.31
Sugar	0.25	0.28	0.37
Cocoa	0.09	0.10	0.12
Lumber	0.02	0.02	-0.01
Orange Juice	0.03	0.03	0.00
Feeder cattle	-0.01	-0.03	-0.10
Lean hogs	0.03*	0.02	0.00
Live cattle	0.04	0.04	0.02
Pork bellies	-0.03	-0.11	-0.41
Gold	0.26	0.28	0.36
Silver	0.43	0.48	0.64
Copper	0.35	0.38	0.52
Platinum	0.16	0.18	0.23

Table 10: Average Realized Betas - Oil only

	S&P GSCI	DJ-UBS CI	Th.Reuters CI
Commodity	$\Delta\beta$	$\Delta\beta$	$\Delta\beta$
WTI crude oil	3.30	1.51	1.03
Corn	0.60	0.66	0.57
Soybeans	0.54	0.40	0.21
Chicago wheat	0.57	0.37	0.34
Soybean oil	0.52	0.26	0.20
Soybean meal	0.42	0.10	0.09*
Rough rice	-0.01	-0.10	-0.12
Oats	0.09	-0.44	-0.46
Coffee	0.29	0.42	0.41
Cotton	0.29	0.36	0.33
Sugar	0.46	0.61	0.54
Cocoa	0.11	0.19	0.17*
Lumber	0.06	0.02	0.00
Orange Juice	0.04	0.15	0.05
Feeder cattle	-0.04	-0.16	-0.20
Lean hogs	0.04	-0.10	-0.12
Live cattle	0.06	0.03	-0.02
Pork bellies	-0.13	-0.60	-0.93
Gold	0.35	0.43	0.39
Silver	0.58	0.77	0.71
Copper	0.49	0.66	0.59
Platinum	0.27	0.33	0.31

Average realized betas pre and post chosen data and difference. Positive significant differences from zero at the 10% and 5% levels in two-sided tests are denoted by * and **, respectively. Light grey cells denote that the difference is not significant at 10% level or higher; dark grey cells denote negative significant differences at 10% level. Remaining cell denote difference significant at 1% level. Commodities in bold are those not included in any index.

4 Conclusions

Over the last decade, commodity futures have become a popular asset class for portfolio investors, as investors increasingly sought out commodities for portfolio diversification after the equity market crash in 2000. This process is sometimes referred to as the financialization of commodity markets. The increase in commodity investment (particularly in the form of investment in commodity indices) coincided with increasing comovement between the returns of different commodity classes and led to speculation about the potential negative effects of commodity financialization.

This paper seeks to add to the literature on commodity financialization and its potential economic implications through an analysis of the underlying drivers of the observed increase in commodity comovement. We build on the work of Barberis et al. (2005) and borrow the concept of 'index inclusion' to describe the financialization process. This allows us to distinguish between two views of return comovement: that comovement in returns is driven by news related to the commodity fundamentals (the traditional explanation for comovement), and that excess comovement can be attributed to frictions or traders' sentiment (the alternative theory of comovement tested for here). We study the comovements between the three main commodity indexes and commodities included and not included in the indexes. We utilise both univariate and bivariate regressions, examining two time periods: 1998-2005 (pre-financialization) and 2005-2011 (post-financialization).

In the univariate regressions, we observe a significant increase in comovement between index non-energy commodities and index starting in 2005. Conversely, for off-index commodities, either no change or a significant decrease in comovement is observed. These results were corroborated through the bivariate regressions: again index commodities exhibit an increase in comovement with the index while off-index commodities do not exhibit increased comovement with the index. We interpret these results as evidence of increased financialization post-2005. However, there are alternative explanations for the regression findings that should be considered. These include non-trading effects (i.e. that the comovement results could have a spurious upward bias due to the greater liquidity and increased trading activity of indexed commodities) and characteristic effects (i.e. that all commodities included in the indices share a common characteristic and that characteristic increases with time). We explicitly test for both of these effects and find that their impact is likely to be minimal. As an additional check of the robustness of our findings, we re-run the regressions for an earlier (pre-financialization) period and find no evidence of an increase in comovement between index or non-index commodities and the indices, which provides further support for the finacialization interpretation.

Finally, to further corroborate the regression findings, we extend our analysis to account for high-frequency returns dynamics by means of the so called realised beta and reach similar results. Our findings cannot be explained by the fundamentals-based view which considers the commodity price bubble and crash as solely explained as driven by fundamentals. We therefore provide new evidence supporting the friction or sentiment based explanations, which we interpret as evidence of the impact of commodity financialization on return comovement.

References

Adams, Z. and Gluck, T. (2015). Financialization in Commodity Markets: A Passing Trend or the New Normal? *Journal of Banking and Finance*, 60:93–111.

Andersen, T., Bollerslev, T., Diebold, F., and Wu, G. (2006). Realized beta: persistence and predictability. In Fomby, T. and Terrel, D., editors, *Advances in Econometrics: Econometric analysis of economic and financial time series.*, volume 20, pages 1–37. Elsevier Sciences: Amsterdam.

Ankrim, E. M. and Hensel, C. R. (1993). Commodities in asset allocation: A real-asset alternative to real estate? *Financial Analysts Journal*, 49(3):20–29.

Barberis, N. and Shleifer, A. (2003). Style investing. *Journal of Finacial Economics*, 68:161–199.

Barberis, N., Shleifer, A., and Wurgler, J. (2005). Comovement. *Journal of Financial Economics*, 75:283–317.

Barndorff-Nielsen, O., Hansen, P., Lunde, A., and Shephard, N. (2011). Multivariate realised kernels: consistent positive semi-definite estimators of the covariation of equity prices with noise and non-synchronous trading. *Journal of Econometrics*, 162:149–169.

Barndorff-Nielsen, O. and Shephard, N. (2004). Econometric analysis of realized covariation: high-frequency based covariance, regressions, and correlation in financial economics. *Econometrica*, 72(3):885–925.

Basak, S. and Pavlova, A. (2013). Asset prices and institutional investors. *American Economic Review*, 103(5):1728–1758.

Basak, S. and Pavlova, A. (2014). A model of financialization of commodities. *working paper*.

Bollerslev, T. and Zhang, B. (2003). Measuring and modeling systematic risk in factor pricing models using high frequency data. *Journal of Empirical Finance*, 10(5):533–558.

Bruno, V., Buyuksahin, B., and Robe, M. (2015). The financialization of food? *Working Paper*.

Buyuksahin, B. and Robe, M. (2014). Speculation, commodities and cross-market linkages. *Journal of International Money and Finance*, 42:38–70.

Campbell, R. (2008). Art as a financial investment. *The Journal of Alternative Investments*, 10(4):64–81.

Cheng, I., Kirilenko, A., and Xiong, W. (2014). Convective Risk Flows in Commodity Futures Markets. *Review of Finance*, 1:1–49.

Cheng, I.-H. and Xiong, W. (2014). Financialization of commodity markets. *Annual Review of Financial Economics*, 6(1):419–441.

Daskalakis, C., Kostakis, A., and Skiadopoulos, G. (2014). Are there Common Factors in Commodity Futures Returns? *Journal of Banking and Finance*, 40:346–363.

Dimson, E. and Spaenjers, C. (2011). Ex post: The investment performance of collectible stamps. *Journal of Financial Economics*, 100(2):443 – 458.

Dovonon, P., Goncalves, S., and Maddahi, N. (2013). Bootstrapping realized multivariate volatility measures. *Journal of Econometrics*, 173:49–65.

Edwards, F. R. and Liew, J. (1999). Hedge funds versus managed futures as asset classes. *The Journal of Derivatives*, 6(4):45–64.

Epps, T. (1979). Comovements in stock prices in the very short run. *Journal of the American Statistical Association*, 74:291–296.

Erb, C. and Harvey, C. (2006). The strategic and tactical values of commodity futures. *Financial Analyst Journal*, 62(2):69–97.

Fattouh, B., Kilian, L., and Mahadeva, L. (2013). The role of speculation in oil markets: What have we learned so far? *The Energy Journal*, 34(3).

Gilbert, C. L. (2010). How to understand high food prices. *Journal of Agricultural Economics*, 61(2):398–425.

Gorton, G. and Rouwenhorst, K. (2006). Facts and fantasies about commodity futures. *Financial Analyst Journal*, 62(2):47–68.

Hamilton, J. D. and Wu, J. C. (2014). Effects of index-fund investing on commodity futures prices. Technical Report NBER Working Paper No. 19892, National Bureau of Economic Research.

Hansen, P., Lunde, A., and Vove, V. (2014). Realized beta garch: a multivariate garch model with realized measures of volatility. *Journal of Applied Econometrics*, 29:774–799.

Irwin, S., Sanders, D., and Merrin, R. (2009). Devil or angel? the role of speculation in the recent commodity price boom (and bust). *Journal of Agricultural and Applied Economics*, 41:377–391.

Irwin, S. H. and Sanders, D. R. (2012). Testing the masters hypothesis in commodity futures markets. *Energy Economics*, 34(1):256 – 269.

Krugman, P. (2008). More on oil and speculation. *New York Times*.

Masters, M. (2008). Testimony before the committee on homeland security and governamental affairs. *Technical report*.

Nazlioglu, S., Erdem, C., and Soytas, U. (2013). Volatility spillovers between oil and agricultural commodity markets. *Energy Economics*, 36:658–665.

Sadorsky, P. (2014). Modeling volatility and correlations between emerging market stock prices and the prices of copper, oil and wheat. *Energy Economics*, 43:72–81.

Stoll, H. and Whaley, R. (2010). Commodity index investing and commodity futures prices. *Journal of Applied Finance*, 1:1–40.

Tang, K. and Xiong, W. (2012). Index investment and the financialization of commodities. *Financial Analysts Journal*, 68(6):54–74.