Commodity Returns: Lost in Financialization

Fahiz Baba-Yara*[1] and Massimiliano Bondatti[2]

[1]Kelley School of Business
[2]Nova School of Business and Economics

January 15, 2023

[First draft: February 2022]

[Preliminary and Incomplete]

Abstract

The flow of investment capital into the commodity futures market dramatically increased around 2004, and this event is referred to as the financialization of commodity markets. We study how this phenomenon has affected average returns in this asset class by examining how returns to popular commodity futures strategies have evolved. We find that about 80% of commodity futures strategies that earned statistically significant average returns pre-financialization have earned an average of a zero return since. Using a six latent factor asset pricing model, we show that this deterioration in strategy returns can be wholly explained by an adverse change in the average returns to systematically priced variation in the cross-section of commodity futures. In robustness tests, we show that the publication of commodity strategies in the academic literature can only explain about 25% of the 51 bps per month reduction that commodity futures strategies have experienced post-financialization.

Keywords: Commodity futures markets, Commodity risk premium, Latent factor models, Financialization of commodities, Commodity investment strategies

JEL Classification: G11, G12, G13, G14, G29

*Baba-Yara: Kelly School of Business, Indiana University, 1309 E. 10th St., Bloomington, USA; fababa@iu.edu. Bondatti: Nova School of Business and Economics, R. Holanda 1, Carcavelos, Portugal; massimiliano.bondatti@novasbe.pt. The authors thank Daniele d'Arienzo, Andrei S. Gonçalves (discussant), Nicholas Hirschey, Francesco Mazzola, Giorgio Ottonello, Melissa Prado and Alberto Rossi for helpful discussions and suggestions. Useful comments from participants to the 2022 Australasian Finance & Banking Conference, and 2022 Nova Finance PhD Pitch Perfect are gratefully acknowledged. Massimiliano Bondatti acknowledges financial support from Fundação para a Ciência e a Tecnologia (FCT). This paper was earlier circulated with the title "Depressed Risk Premia or Mispricing; Where Did the Commodity Returns Go After Financialization?". All errors remain our own.

Introduction

Starting in 2004, a significant amount of investment capital began flowing into commodity futures as retail and institutional investors sought to diversify their portfolios beyond traditional asset classes like stocks and bonds (Basak and Pavlova (2016)). This led to an increase in investment capital from about $15 billion to around $200 billion between 2003 and 2008, with most of it being invested in net long positions (CFTC (2008), Henderson et al. (2015)). The primary vehicle used to gain exposure to these markets was commodity tracking indexes such as the S&P Goldman Sachs Commodity Index (henceforth, SPGCI) and the Dow Jones Commodity Index (henceforth, DJCI), among others.

This phenomena is known as the financialization of commodity futures markets and has been studied extensively by policy makers and the academic literature. Previous research has primarily focused on how the influx of index capital affects commodity prices and volatility. Additionally, the literature has also examined how the change in price informativeness of commodities affects firms with significant economic exposure to commodities due to their production process. Specifically, Brogaard et al. (2019) show that financialization of commodity markets has hindered the ability of these firms to extract signals from market prices to aid their decision making.

In this paper, we ask whether or not the average returns of commodity futures strategies have been significantly impacted by the financialization and influx of index capital.

The existing literature on commodity futures strategies suggests that the demand and supply imbalance is the fundamental driver of commodity strategies' positive average returns. More specifically, the strategies inform speculative capital whether extra capital is needed on a commodity's demand or supply side to help futures markets clear. If this is indeed the case, then it is relevant to ask if the influx of commodity index traders, who are predominantly net long, have left the profitability of commodity strategies unaffected. Theoretical research has shown that financialization affects commodity spot and futures prices (see Keynes (1930), and Basak and Pavlova (2016)) and future prices informativeness (see Goldstein and Yang (2021)). Empirical studies such as Singleton (2014), Henderson et al. (2015) and Brogaard et al. (2019) find that the index flows have led to significant changes in commodity futures prices, informativeness, and volatility. As these prices and volatility are key factors that inform commodity futures strategies, it is essential to investigate their performance post-financialization. If commodity futures prices are less informative for business decision-making as Cheng et al. (2015) and Brogaard et al. (2019)) have shown then, they may similarly be less informative for commodity futures strategies.

In this paper, we examine twenty different commodity futures strategies. Our findings support the assertion that the informativeness of commodity prices has significantly decreased after the financialization of commodity markets. Before 2004, thirteen of the twenty strategies we examine produce at least marginally significant returns. However, after financialization, this number drops to only three strategies.

We propose and test two potential explanations for this stylized fact. The first hypothesis posits that many of these strategies were previously able to exploit idiosyncratic mispricings across different commodity futures. With the influx of investment capital following financialization, these mispricings have been competed away. The second hypothesis suggests that a limited number of systematic commodity factors, such as risk or behavioral factors (see Kozak et al. (2018)), drive returns in these strategies. If this is the case, an influx of index investors may depress the average returns of these factors and, subsequently, the returns to commodity futures strategies.

To this end, we propose a six-factor asset pricing model and show that this model explains most of the priced variations to all the twenty commodity futures strategies. We also show that the average return of the individual factors has undergone substantial decay after the influx of investor capital after 2004. And so the deterioration in returns to commodity futures strategies derives from a reduction in returns to the handful of factors that drive returns in this market.

Following the existing literature, we examine the potential relation between the pricing factors in the six-factor model and the macro-economy. We focus on the systematic factors that have undergone changes around the financialization period and explain a significant portion of the variation in commodity returns. Overall, we uncover several relevant channels through which economic shocks relate to commodity prices. Specifically, we find evidence that shocks to inflation, commodity volatility, and, to a lesser extent, interest rate and FX-related factors drive variations in the systematic latent factors in commodity futures markets. Additionally, similar to previous research (Bakshi et al. (2019)), we also find that global equity volatility is closely related to the latent factors in the commodity market.

To address the concern that the depression in strategy returns might be attributable to the popularization of these strategies instead of the financialization of the market, we run tests similar to those of McLean and Pontiff (2016). We find that about 80% of the average decline in average returns is attributable to the financialization of commodity markets and the remaining 20% coming from the popularization of these strategies.

The results in this paper are important for a number of reasons. First, we identify a new channel through which financialization of commodity futures has affected the asset class. We particularly show that the influx of investment capital has depressed returns to most commodity strategies. This is a previously unexplored channel. Second we propose a simple agnostic linear asset pricing model that summarises the cross-section of commodity futures strategies. This model shows that returns to commodity futures strategies exclusively come from exposure to a handful of systematic factors in this cross-section. Finally, we show that the deterioration in strategy returns can be tied to significant exposure to commodities in popular long only commodity indexes such as the DJCI.

This work contributes to three main strands of the literature. First, it contributes to the study of cross-sectional and time-series predictability in commodity futures markets. A number of recent studies have proposed several variables that successfully predict variation in commodity returns: carry and basis (Szymanowska et al. (2014), Bakshi et al. (2019), Koijen et al. (2018)), momentum (Miffre and Rallis (2007), Szymanowska et al. (2014) and Bakshi et al. (2019)), basis-momentum (Boons and Prado (2019)), reversal (Bianchi et al. (2015)), value (Asness et al. (2013) and Baba Yara et al. (2021)), coefficient of variation (Dhume (2010)), volatility and inventory (Gorton et al. (2013)), open interest (Hong and Yogo (2012)), hedging pressure (De Roon et al. (2000), Basu and Miffre (2013) and Kang et al. (2020)), liquidity (Marshall et al. (2012) and Marshall et al. (2013)), inflation and the dollar (Erb and Harvey (2006) and Gorton and Rouwenhorst (2006)), skewness (Fernandez-Perez et al. (2018)) and level (Bakshi et al. (2019))[1]. We contribute to this strand of the literature by comparing the pre-and post-financialization average returns of these strategies. Additionally, we decompose the returns into what can be explained by systematic variation in the cross-section of commodity markets and what fraction is idiosyncratic to a particular commodity. We find that using a six factor linear asset pricing model, the returns to all the twenty strategies we study are attributable to systematic variation in this cross-section.

The second strand of the literature we contribute to studies how the financialization of commodity markets has affected the cross-section of commodity futures. CFTC (2008) finds that open interest and investment inflow into commodity indexes increased substantially around 2004. Cheng and Xiong (2014) discuss the different ways through which this phenomena can affect the commodity futures and subsequently the real economy. Boons et al. (2014), Büyükşahin and Robe (2014), Christoffersen and Pan (2018) and Melone et al. (2021), among others, find that the correlation between stock and commodity futures markets dramatically changed around this point. Brogaard et al. (2019) show that index investing, which has increased during the post-financialization period, reduces price informativeness of index commodities and, consequently, decreases the sensitivity of index commodity firms to commodity futures prices. Closely related, Hamilton and Wu (2014) find a change in oil futures risk premia since 2005 related to the increasing importance of index-fund investing relative to commercial hedging in affecting crude oil futures risk premia. While, theoretically, Basak and Pavlova (2016) and Goldstein and Yang (2021) argue that commodity futures prices, volatilities, price informativeness and correlations across commodities and with other assets (e.g. stocks) increase with the financialization. Lastly, Baker (2021) calibrates a macro-finance model for storable commodities. The

[1]Section 1.2 and the Appendix (A.2) discuss in depth the economic rationale behind each variables and how they are constructed.

author finds a decrease in the risk premium of storable commodities in response to the financialization. We contribute to this literature by showing that returns to a number of prominent commodity futures strategies have materially decayed towards zero after the financialization of commodity markets. We further show that the fall in strategy returns has happened mainly because the average returns to the pricing factors in this asset class have significantly fallen following financialization.

Finally, our paper contributes to the asset pricing literature that studies return predictors after publication. The seminal work in this literature (McLean and Pontiff (2016)) studies the post-publication stock return predictability of 97 variables and find that returns are on average 58% lower. Similar results hold also in forex (Bartram et al. (2022)). Hou et al. (2020) find that most equity anomalies fail to hold up to currently acceptable standards for empirical finance. However, a Bayesian modeling framework Jensen et al. (2021) show that a majority of anomalies replicate. We focus on commodity markets as against equity and forex markets. Additionally, we do the exercise in the spirit of McLean and Pontiff (2016) and Bartram et al. (2022), but through the lens of the financialization of a market. We find that the commodity futures market has also experienced a significant decay in returns in recent years. However, we find that only about 25% of this decay can be attributed to the popularization of the strategies through publication compared to financialization.

The reminder of the paper is organized as follows. Section 1 presents the data and commodity strategies. Section 2 examines how the financialization of commodity markets has affected average returns to commodity strategies. Section 3 explains the economic rationale behind the decay and section 4 provides robustness. Section 5 explore an index-investing mechanism that can help in rationalizing our findings. Finally, section 6 concludes.

1 Data and characteristic construction

Our sample starts in March 1986 and ends in August 2021. We retrieve end-of-day data on liquid commodity futures contracts from the Commodity Research Bureau (CRB) from March 1986 to December 2014. We extend the sample to August 2021 with data from Datastream (now Refinitiv) and Factset. We pick this starting date to have futures returns data for a reasonable number of commodity strategies in the cross-section while balancing the pre- and post-financialization periods[2]. Overall, we study 32 commodity futures contracts belonging to four major sectors: agriculture, livestock, energy, and metal.

1.1 Commodity futures returns

We conduct most of our analysis at the monthly frequency and compute holding period returns using end-of-month prices. More specifically, we follow Bakshi et al. (2019) in the construction of commodity excess returns between period-t and $t+1$. At the end of each month-t, we enter a position in the commodity-specific futures contract with the second shortest maturity while guaranteeing that its first notice day is *after* the end of month $t+1$. By rolling into the shortest maturity contract before the first day of notice for each commodity, we guarantee that we are never forced to take physical delivery of a commodity[3]. This convention is broadly consistent with Hong and Yogo (2012), Gorton et al. (2013), among others.

We compute the returns of the long and short commodity futures positions as:

$$
\begin{aligned}
r_{t+1}^{long} &= r_t^f + \frac{1}{F_t^{(1)}}(F_{t+1}^{(1)} - F_t^{(1)}) \\
r_{t+1}^{short} &= r_t^f - \frac{1}{F_t^{(1)}}(F_{t+1}^{(1)} - F_t^{(1)})
\end{aligned}
$$

where r_t^f is the interest earned on a fully collateralized futures position, and $F_t^{(1)}$ is the

[2] At each point in time we have available at least 25 commodities to be allocated into (at most 5) portfolios for each sorting variable. Moreoever, the starting date is aligned to Szymanowska et al. (2014).

[3] We refer the interested reader to Bakshi et al. (2019) for further details on the futures return construction, and to their Table I in the internet Appendix for details on the first notice day convention.

price of the next maturity futures contract at the end of month t [4]. Excess returns between period t and $t + 1$ are then calculated as:

$$e_{t+1}^{long} \equiv r_{t+1}^{long} - r_t^f \tag{1}$$

$$e_{t+1}^{short} \equiv r_{t+1}^{short} - r_t^f$$

Table A.1 in the Internet Appendix provides the descriptive statistics for the individual commodity futures excess returns. For about 25 out of the 32 commodities we study, the average returns is positive. This means that a long-only rolling strategy as described above will typically generate a positive return over our sample period. As has been noted previously in the literature (see Bakshi et al. (2019)), the volatility of individual commodity strategies tend to be very high and so the resulting Sharpe ratios of these long-only rolling strategies are typically below 0.25. Commodity futures also tend to be positively skewed as we observe in Table A.1, which provides suggestive evidence for why it is the preferred asset class of Trend Following traders. Lastly, all the commodities are available for at least two-thirds of the sample period.

1.2 Characteristic definitions

Several variables have been shown in the literature to predict variations in the cross-section of commodity futures returns. In this section, we detail how we construct these characteristics and subsequently form portfolios by sorting the 32 commodities we study on these characteristics. For each characteristic, we follow the portfolio sorting scheme as proposed in the article in which the predictor was introduced.

First, we construct the commodity carry strategy as in (Bakshi et al. (2019) and Koijen et al. (2018)) and the basis strategy as in (Szymanowska et al. (2014) and Boons and Prado (2019)). Both predictors have been shown to predict returns in the cross-section and in the time series of commodity futures.

[4]Consequently, $F_{t+1}^{(1)}$ is the price of the next maturity futures contract observed at the end of month $t+1$ (i.e., when the position is closed), and $F_t^{(0)}$ is the price of the front-month futures contract observed at the end of month t (i.e., when the position is open).

Second, we construct different versions of long- and short-term momentum following Miffre and Rallis (2007), Szymanowska et al. (2014), Boons and Prado (2019) and Bakshi et al. (2019). Specifically, we construct two 12-month momentum strategies (*Mom12* and *MoB12*), one 6-month (*Mom06*), 3-month (*Mom03*), and 1-month momentum (*Mom01*). Whereas *MoB12* follows the description in Boons and Prado (2019), *Mom12* follows Szymanowska et al. (2014). Additionally, we include a reversal factor (*Rever*) as in Bianchi et al. (2015).

Thirdly, we construct the basis-momentum factor (*BaMom*) of Boons and Prado (2019) that leverages both momentum and basis fundamental signals. This factor is related to the slope and curvature of the commodity futures curve; and it is consistent with imbalances in supply and demand of future contracts that appear when the market-clearing ability of speculators and intermediaries is impaired.

Fourth, we consider the set of volatility based measures. The coefficient of variations computed using spot prices (*CVDhu*) comes from Dhume (2010), while the one using returns (*CVSzy*) comes from Szymanowska et al. (2014)). We also include a volatility factor (*Volat*) constructed as in Gorton et al. (2013), which the authors show to be related to inventory.

Fifth, we construct the inventory (*Inven*) predictor, inspired to Gorton et al. (2013)[5]. The theory of storage (Kaldor (1939), Working (1949), Brennan (1958) and Deaton and Laroque (1996)) relates to the timing option that is intrinsic when holding a storable commodity. The action of delaying the consumption of a commodity to tomorrow (i.e. when its supply might be scarce) pushes its prices above the value of consuming it today and produces the convenience yield of holding the commodity. Hence, in this theory, the carry of a storable commodity directly maps into the cost of storing it. Holders of inventories earn a convenience yield that is a declining and convex function of inventory. Thus, the commodity futures risk premium decreases with inventories. We refer also to Cheng and Xiong (2014) for an in depth review of this topic.

The sixth set of variables are related to the hedging pressure theory of Keynes (1930)

[5]We thank Martijn Boons for kindly sharing the inventory data with us. The inventory factor covers the period only up to 2011.

and Hicks (1939). Hedgers in the futures market tend to be in a net short position and want to get rid of their risk. To incentivize other market participants to take on this risk, the hedges need to offer their counter-parties a premium to induce them to take the long position. Hedging pressure (*HedPr*) has been shown to affect commodity excess returns by Bessembinder (1992), De Roon et al. (2000) and Basu and Miffre (2013), among others. We include a related predictor, open interest (*OpeIn*), which Hong and Yogo (2012) have show to similarly predict commodity prices beyond imbalances among hedgers.

The seventh set of variables seek to exploit the cross-sectional heterogeneity in the conditional correlation between commodities returns and prominent macro-variables (see Erb and Harvey (2006), Gorton and Rouwenhorst (2006) and Szymanowska et al. (2014)). We include an inflation-β (*InflB*) variable, which is based on the conditional correlation between commodity future returns and inflation; and a dollar-β (*DollB*) variable, that seeks to exploit a similar correlation between commodity futures returns and exchange risk (commodity future prices are expressed in a currency, typically the U.S. dollar).

Eight, we sort portfolios on a liquidity variable (*Liqui*), namely the Amivest measure of Amihud et al. (1997), which is inspired by Marshall et al. (2012) and Marshall et al. (2013). Expected commodity excess returns can indeed reflect the liquidity of the contract since liquidity is potentially different across futures or maturities.

Ninth, we construct a standard value factor (*Value*), following Asness et al. (2013) and Baba Yara et al. (2021). The value factor is computed using long-term past returns which rests on the long standing literature that finds that past returns and book-to-market ratios are correlated (see De Bondt and Thaler (1985), Daniel et al. (1998), and Gerakos and Linnainmaa (2018)).

Tenth, we construct a skewness (*Skewn*) factor following Fernandez-Perez et al. (2018). As the authors show, in line with predictions of theories on investors' skewness preferences or on selective hedging, this factor delivers significant returns and is able to explain the cross-section of commodity futures returns.

Finally, we include a level factor as in Bakshi et al. (2019). This factor is the equally-weighted commodity strategy (*Averg*) that goes long in all the commodities available at month t.

2 Financialization and the cross-section of commodity futures returns

In this section, we provide a brief background on the financialization of commodity futures markets and highlight how that has affected a number of commodity futures strategies.

2.1 Background on financialization

Traditionally, the commodity futures market has been dominated by two main participants, commercial hedgers and noncommercial traders. Commercial hedgers are primary producers of commodities such as farmers, and primary users of commodities such as oil refineries who seek to hedge their business activities against spot-price risk. Noncommercial traders are managed money traders such as hedge funds and Commodity Trading Advisors (CTAs) who take a position on the long or short side to help balance out demand and supply between primary commodity producers and users.

Around the turn of the millennium, fund flows into commodity futures as an asset class began to grow astronomically. Investment inflows into this space grew from about \$20 billion in 2003 to more than \$200 billion in 2008 and to about \$300 billion in 2010 (CFTC (2008) and Irwin and Sanders (2011)). The total U.S. exchange-traded futures and futures option trading volume moved instead from about 630m contracts per year in 1998 to about 3.2b contract p.y. in 2007, with the growth spread across all the commodities. As shown by Brogaard et al. (2019), open interest across a number of commodities were relatively flat between 2000 and 2003, but increased tremendously after 2004. At the same time, there was also a significant change of mix of market participants with the strong entrance of institutional and index investors (Boons et al. (2014), Irwin and Sanders (2011), Brogaard et al. (2019)). As of 2008, the total net notional value of funds invested in commodity indexes was held by "Index Funds" for approximately 24% and by "Institutional Investors" for approximately 42% (CFTC (2008)). Overall, this change in market structure has been dated to have happened around 2004 (see Boons et al. (2014), Basak and Pavlova (2016) and Brogaard et al. (2019), among others) and is referred to as the financialization of commodity futures markets in the literature.

The exact effect of this kind of financialization is still being studied by academics, regulators and practitioners. Theoretically, Basak and Pavlova (2016) and Goldstein and Yang (2021) show that the rapid increase of indexers should affect commodity prices and volatility. Empirically, the evidence is murky. Stoll and Whaley (2010) and Hamilton and Wu (2015) both find no evidence of the increased index flow affecting commodity prices or volatility. On the other-hand, Singleton (2014), Henderson et al. (2015) and Brogaard et al. (2019) find evidence to the contrary.

Additionally, the financialization affected the informativeness of commodity prices. Goldstein and Yang (2021) show that, as the financialization becomes larger in size, the noise brought in the commodities market by financial hedgers prevails and overcomes the positive effect of financial speculators on price efficiency, thus reducing the overall price informativeness. Similarly, Brogaard et al. (2019) find a decrease in commodity price informativeness in response to the commodity financialization. However, they attribute this effect to the index investing mechanism. An increasing number of empirical studies in other asset classes shows indeed that index investing, in general, leads to worse price informativeness (and to higher price volatility), e.g. Israeli et al. (2017), Ben-David et al. (2018) and Coles et al. (2022).

Our paper contributes to the debate above by introducing a new channel through which the increased indexing in this market has affected the dynamics therein. Whereas the traditional debate has been on how financialization has affected price and volatility, our contribution focuses on the returns to commodity returns which one can think of as compensation for speculators. As commodity prices become less informative in response to the commodity financialization (Brogaard et al. (2019) and Goldstein and Yang (2021)), the signals on which most of the investment strategies are built plausibly lose their overall informativeness as well, thus compromising the profitability of the strategies. As far as we are aware, we are the first to study the implications of the commodity financialization for the commodity futures trading strategies. This is an instructive exercise as most of the indexers who have flooded this market since 2004 have done so on the belief that commodity futures guarantee a risk premia which one should gain unconditional exposure to.

2.2 Disentangling commodity futures portfolio returns

As is standard in the literature, we sort commodities on the characteristics described in section (1.2) and then form long-short portfolios from the resulting extreme portfolios. The number of portfolios formed from the sorts and the definition of the long and short legs of each strategy follow from the studies referenced above[6].

Table 1 reports the summary statistics for all the long-short commodity futures strategies we study. Over the entire sample, about 50% of the strategies have an average return that is statistically significant at the 5% level. An additional 15% are significant at the 10% level[7]. Similar to what we observed for individual commodity futures, the long-short portfolios have similar volatilities. In contrast to the portfolio underlying, the Sharpe ratios are much higher. Whereas Sharpe ratios of less than 0.25 are common place for individual commodities, Sharpe ratios for strategies that have at least a marginally significant return are higher than 0.25.

When we decompose the return of the strategies into their pre- and post-financialization component, we observe a striking result: of all the strategies with a significant average return over the full sample, only Carry and Skewness remain significant in the post-financialization period[8]. Put differently, about 80% of the commodity strategies that are significant over the entire sample, seem to have generated their average return from a period before the current regime[9].

For most of these strategies, the loss of statistical significance is not coming from an

[6]Section A.2 in the Internet Appendix provides further details on how we construct each characteristic and form long-short portfolios from the sorts.

[7]Figure A.1 in the Appendix summarizes these results.

[8]Figure A.2 in the Appendix summarizes these results.

[9]When we move instead the split date even just 5 years before the commodities markets financialized (i.e. in January-1999), there are almost as many strategies that returns average returns different from zero in the periods before and after the split date (specifically, ten vs seven, see Table A.2).

increase in volatility but rather a reduction in average returns. This is an interesting fact to document as the discussion in the literature about how financialization has affected the commodity markets has mainly focused on the price and volatility channels. Our results here show that volatility of commodity futures strategies have remained relatively unchanged, however the average returns they generate have significantly deteriorated. It is also worth pointing out that hedging pressure is the only strategy with an insignificant average return in the pre-financialization period but a significant premia post financialization. This is interesting because hedging pressure is a strategy with a strong theoretical foundation for why it should earn a risk premia in this asset class.

Overall, this table provides a novel empirical finding in the commodity futures literature: the returns to several commodity futures strategies fade away post-financialization. Hence, we show that the financialization of commodity futures markets, which saw the flood of investment capital into this asset class, has affected the market through a channel which has until now gone unexplored[10]. The sudden influx of passive investment capital into an asset can significantly depress the average returns.

This financialization effect is also intuitively shown in Figure 1, where we plot the difference in returns to the strategies post-financialization against the returns pre-financialization. The negative effect exists for most of the strategies individually, and the commodity anomalies with higher returns pre-financialization tend to show larger declines in returns post-financialization.

3 What explains the decay in average returns?

In this section, we propose and test two alternative hypothesis that can help explain the decay in average returns to the commodity futures strategies returns.

[10] A decay in returns is observed in other asset classes (namely, equity and forex) after the publication of the trading strategy in the academic literature. Hence, one might worry that the decay in the returns to the commodity strategy we observe post-financialization might actually be capturing a similar phenomenon. In Section 4.1 we run robustness tests that rule out this potential alternative story.

3.1 Systematic or Idiosyncratic Deterioration?

We have shown previously that the average return to more than 80% of the profitable commodity strategies have decayed since financialization. There are two main ways this could have happened. First, the returns that the factors were generating were purely idiosyncratic mispricing and so once more investment capital came into the asset class, that mispricing has been competed away. The second hypothesis takes seriously the fact that a simple linear factor model can price the cross-section of these commodity futures strategies. If this is the case, then the reduction in returns to the strategies has to come from more fundamental drivers of variation in the economy. Given the recent work of Kozak et al. (2018), the existence of a handful of pricing factors tells us that the returns to the strategies stem from the small set of primitives. This still leaves the question of whether or not the primitives are driven by risk factors or that they are behavioural. A final hypothesis is that both hypothesis one and hypothesis two partially explain the empirical fact we document.

To disentangle these hypotheses above, we consider a simple linear asset pricing factor model that does not take a stance on what the pricing factors are. Using the novel risk-premium principal component analysis (henceforth, RP-PCA) of Lettau and Pelger (2020), we extract latent factors from the cross-section of commodity futures portfolios.

RP-PCA is a generalization of PCA that seeks to extract latent factors that simultaneously explains variations in the time-series of the portfolios while explaining variations across average returns. As a consequence, the factors estimated from this new method fits both the time-series and the cross-section of expected returns compared to standard PCA where the latent factors attempt to only explain variation in the time-series. We select six factors following the scree plot in Figure 2[11]. To shed light on which hypotheses the evidence favor, we run Fama and MacBeth (1973) tests of the returns to the commodity futures strategies on the first six RP-PCs over the full sample and the two samples centered around financialization (2004)[12].

[11] This plot of the first 15 eigenvalues shows (up to) three strong factors and three weaker, but potentially relevant, factors.

[12] The overall results do not change when considering a seven- or eight-latent factor model, as the

More specifically, this asset pricing method builds on the idea that, absent arbitrage opportunities, the Euler equation implies that risk-adjusted returns on each zero-cost portfolio should be zero on average. With excess returns to the commodity portfolios defined as Rx_t and a stochastic discount factor (SDF) as M_t, the following should hold:

$$E[Rx_t M_t] = 0 \tag{2}$$

More specifically, a SDF that is linear in factors can be expressed as $M_t = 1 - (h_t - \mu_h)'b$, where h_t is the vector of pricing factors, b is a vector of factor loadings, and μ_t is a vector of the factor means. This SDF specification admits a beta representation of the form:

$$E[Rx_{i,t}] = \lambda'\beta_i \tag{3}$$

where the risk premia to some strategy i depends on the price of risk (λ) and how the strategy loadings on the factors (β_i)[13]. Thus, for us, this represent the regression coefficients of the excess returns to each commodity investment strategy on the (latent) risk factors, i.e. the RP-PCs. As mentioned, we estimate these coefficients through a standard Fama and MacBeth (1973) two stage procedures, as common in the literature.

Table 2 reports the results. From panel A, it is evident that the six latent factors can explain all the variation in average returns. As a result the alpha in the second-stage of the test is both statistically and economically indistinguishable from zero. When we focus on both Panels B and C, which represent the pre-financialization and post-financialization periods, we can see that the alpha remains statistically insignificant. This is evidence against the first hypothesis. That is to say the strategies do not appear to represent idiosyncratic mispricing that was arbitraged away after the financialization.

Ruling out hypothesis one, also rules out hypothesis three. When we focus on the risk premia to the different PCs, we can see that over 60% have experienced a reduction. This reduction is over 100% for PC3 to as small as 2% for PC2. PCs 5 and 6 have

additional RP-PCs have little explanatory power in all the sub-samples and their prices of risk are not significant (see Table A.4).

[13]The factor prices and the factor loadings are related through the covariance matrix of the factors (Σ_h): $\lambda = \Sigma_h b$.

both experienced an increase. Taking together, the results provide suggestive evidence in support of hypothesis two. This is to say, the reason why almost all commodity futures strategies haven't generated significant average excess returns after financialization is because the (risk or behavioural) systematic primitives that drive them have adversely changed.

One may be worried that even-though the joint tests show that the latent factor model explains all the variations in the commodity strategies we study, the same might not hold true for individual asset pricing tests. To address this concern, we report the first stage Fama and MacBeth (1973) results in Table A.3 in the internet Appendix. From this, it is evident that the six-latent factor model prices all twenty commodity futures factors in the full sample and both sub-samples.

Figure 3 reports the ten-year rolling average of the latent factors[14]. From this, we can clearly see that average returns of the dominant systematic factor (PC1) continues to decay. Whereas the average return of this factor was consistently above 3% pre-financialization, it is clearly trending towards a negative price of risk. The last ten-year average of this factor is currently negative. Additionally, the figure shows that the third latent factor (PC3) has also experienced a significant deterioration. Finally, we can also observe the level change in the price of risk of PC6. This latent factor has an average return that is marginally below zero pre-financialization and significantly positive in the post-financialization period.

Taken together, the results from the asset pricing test strongly support the hypothesis that the influx of investment capital into the commodity futures market around 2004 has depressed the returns to the strategies we study through the handful of the systematic factors they load on.

3.2 Characterising the latent factors

We have shown that variation in returns to the commodity strategies are driven by a handful of systematic factors. Given that this is a new factor model in the cross-section

[14]Using total open interest as a proxy for flows, Figure A.3 shows that the flows of capital into the commodity futures market kept increasing even after the large sudden spike around the financialization.

of commodities, we take the opportunity to better characterise some of its properties[15].

3.2.1 Factor interpretations

We start by analysing how the individual latent factors are formed from the individual commodity strategies. To this aim, we run a factor-strategy regression exercise to relate the strategy to the factors. The results are presented in Table 3.

The results suggest that the first factor is mostly informed by returns to momentum based strategies. The univariate regressions involving the momentum based strategies have an R-squared of at least 50% and high estimated coefficients. The evidence is not as clear-cut on the second factor; although the momentum strategies still appear to play a role in addition to the volatility strategies. PC3 is instead mainly a reversal factor, but also inflation and dollar-index strategies play a significant role in driving variations in this PC. Eventually, the sixth factor is mostly informed by liquidity, open interest, volatility, and hedging pressure. This factor can therefore be labeled as a possible market friction factor.

3.2.2 What are the macro-financial drivers of the latent factors?

We have established that the six-factor latent factor model spans the space of commodity futures strategy returns. To further shed light on the potential macro-drivers of the latent factors, we run univariate regressions of the individual factors, i.e. the RP-PCs, on a set of macro-financial variables:

$$RPPC_t^j = \phi_0 + \phi_1 \Delta MacroFin_t^i + \epsilon_t \tag{4}$$

where, $RPPC_t^j$ is the time t return of latent factor j from Table 2 extracted as in Lettau and Pelger (2020)), and $\Delta MacroFin$ are the innovations to the macro-financial factors. Test statistics are computed with Newey and West (1987)-corrected standard errors.

[15] As PCs 1, 3 and 6 are the ones that experience a significant change around the financialization and, together with PC 2, explain most of the variation across the different subsamples, we will focus our analysis on the properties of these latent factors. However, we refer the interested reader to Appendix A.8.1 for a broader discussion that includes also an analysis of the other latent factors.

The set of macro-financial variables we consider belongs to those variables that are gaining attention in the asset pricing literature that tries to understand the link between the macroeconomy and the (global) financial conditions on the one hand, and the asset returns on the other[16]. We present the results from this exercise in Table 4.

We find shocks to: i) global equity volatility, ii) commodity volatility and iii) inflation to be important sources of variations for the latent factors. Specifically, global equity volatility significantly drives variations in PC1, as well as PC3 and PC5. Similarly, Bakshi et al. (2019) find that equity volatility also drives variation in the carry pricing factor they include in their commodity futures asset pricing model. However, unlike the same authors, we find evidence that a similar measure of volatility constructed from commodity returns does explain variations in some of the pricing factors, namely PCs 2, 3 and 5. Additionally, we find that shocks to inflation (CPIAUCSL and WPSFD49207) strongly matter across the commodity-latent factors[17]. In particular, they drive a lot of variation in PC3, and contribute to variations in PCs 1, 2, 4 and 5.

Besides these three main macro-financial factors, shocks to other variables are also additional sources of variation that (more weakly) affect the latent factors whose dynamics change around the financialization. Specifically, variations in PC1 are driven also by shocks to the forex factors (TWEXAFE and sliq). While, variations in PC2 partly come from financial variables that can loosely be linked to variations in discount rates.

PC3 is instead negatively related to positive shocks in the S&P 500 (as also PC5) and the global financial cycle, and positively related to the default spread. Taken together, this suggests that PC3 is partly related to investor risk aversion similar to what Bakshi et al. (2019) find for their carry factor.

Lastly, PC6 only loads on industrial production and the forex factor, but with an opposite sign in the loadings with respect to PC1.

Overall, although previous literature finds that only a handful of macro-financial variables are relevant for the cross-section of commodity futures, we find the opposite

[16]Table A.5 in the Appendix lists and describes more in details the variables and the sources from which they are retrieved.

[17]In this respect, it is worth pointing out that the role of inflation risk in driving asset prices is drawing a remarkable attention in the current academic debate (see Fang et al. (2022), among others).

results. For most of the macro-financial variables we study, we find that shocks to them can be traced to at least one of the latent pricing factors.

As we saw in the previous section, the average returns of the latent factors tend to change between pre- and- post-financialization. Table A.6 and Table A.7 in the Appendix repeat the univariate regressions of the macro-financial factors on the RP-PCs, respectively, over the pre- and post- financialization periods. Interestingly, it appears that variations in macro-financial risks tend to be a more relevant determinant of variation in the latent risk factors since the occurrence of the financialization.

4 Robustness tests

In this section, we test and rule out alternative explanations for the decay in returns to the strategies we study.

4.1 Pre- vs Post-Financialization, and Academic Research

We have so far argued that financialization is the main driver of the decay of returns to commodity strategies. One alternative story to this is that the publication of the strategies in academic journals is indeed the event that triggers the decay and not necessarily financialization. This is a compelling alternative because most of the strategies we study were published post-financialization. To study this question, we undertake a test similar in spirit to McLean and Pontiff (2016) in equity and Bartram et al. (2022) in forex. Formally, we run variations of the following baseline specification to test this alternative channel:

$$R_{i,t} = \alpha_i + \beta_1 \textit{Post-Financialization Dummy}_{i,t} + \beta_2 \textit{Post-Publication Dummy}_{i,t} + e_{i,t}. \quad (5)$$

Specifically, the first test we run is to regress the returns to all the strategies we study on a dummy that takes the value zero before January 2004 (pre-financialization) and one afterwards (post-financialization). Table 5 reports the results. Standard errors for all the tests are clustered on time. The coefficient on the dummy for this specification is negative and statistically significant which confirms the results from our portfolio sorts. Starting from around 2004, the average returns to a typical commodity futures strategy has decayed about 51 bps per month. The pre-financialization average returns of the strategies is about 78 bps per month.

The second test we undertake is to regress the returns to the commodity strategies on a dummy that takes the value zero before the publication date of a strategy and one afterwards. The coefficient from this second specification is similarly negative and significant. The result suggests that the typical commodity strategy has experienced about a 46 bps per month decay in returns after the strategy was published in an academic journal.

Given that most of the commodity futures strategies we study were published after financialization of the market, it is important to disentangle which of these two alternatives drives what fraction of the decay. To this end, we regress the returns to the strategies on two dummies. The first dummy takes the value of zero before financialization and one afterwards. The second takes the value of zero before publication and one afterwards. We report the results as specification 3 in Table 5. As is evident, the post-financialization dummy subsumes the post-publication dummy. After financialization, the average commodity strategy has lost about 44 bps per month on average. After publication, the strategies lost an additional 14 bps on average although this estimate is not significant. The results therefore show that it is the financialization of commodity markets that is driving the most of the decay in returns and not the publication of the strategies.

To verify the robustness of this conclusion, we run a fourth specification where the post-financialization dummy takes a value of zero before financialization and after publication and a one in-between. The publication dummy takes a value of one after publication. This allows us to strongly isolate the financialization effect. The results in specification 4 show confirm the results in 3 for the financialization effect[18].

[18]Additionally, if we regress the returns on the stock market on the financialization dummy, the estimated coefficient is small but positive. While, including the returns to the market as control in the specifications in Table 5 does not affect the results. This evidence is suggestive that the behaviour of the overall market does not seem to explain our findings. The results to these additional tests are available

Overall, around 75% of the observed decay in returns to commodity strategies can be attributable to the financialization of the market and the rest comes from the publication of the strategies.

5 Index flow mechanism

We have shown that the deterioration in average returns of commodity futures strategies post-financialization is a strong stylized fact in the data. In this section, we provide evidence in favour of the index channel as being the potential driver of the deterioration we observe.

Data on constituents of the Dow Jones Commodity Index and its weights come from Standard & Poor's, and are available from January 2000.

To test this potential mechanism we run the following baseline specification:

$$R_{i,t+1} = \alpha_i + \beta_1 D_{i,t} + \Gamma D_{i,t} * \delta_t + e_{i,t+1}. \tag{6}$$

where: i) $R_{i,t+1}$ are the returns to commodity investment strategy i, ii) α_i is a dummy capturing strategy fixed effects, iii) $D_{i,t}$ is dummy that takes value one if the strategy i at time t has exposure to any commodity in the top-n weighted commodities in the DJCI index (with $n = \{1, 3\}$)[19], and iv) $D_{i,t} * \delta_t$ is a dummy that at time t is equal to one for all strategies that have exposure to a commodity whose weight is in the top-3 of the index [20]. This last term, which can be interpreted as an "exposure by time fixed effects" indicator, allows us to control for the increasing inflows of capitals in the commodity markets (and, so, in the commodities index) over time. This variable also absorbs variation coming from other potential time-varying confounding omitted effects that affect index exposed and non-index exposed strategies differently.

Results to these tests are reported in Table 6. β_1 is our coefficient of interest, as

upon request.

[19]The top-3 weighted commodities at each point in time, alone, account on average for around 40% of the overall DJCI weights. The weights of commodities in the index precipitously fall off such as the highest weighted commodity at each point in time is, on average, almost 4 times larger than the average weight of the third highest weighted commodity.

[20]Γ is a vector of coefficients: $\Gamma = \gamma_1, \ldots, \gamma_t, \ldots, \gamma_T$, with $t = 1, \ldots, T$.

it captures the average effect on the returns to the strategies of trading on commodities with high weights in the commodity index. As we observe in column (1) and (2), having exposure to the index has a significant negative effect on the average returns of a strategy. In other words, a commodity investment strategy sees about a 70 bps drop in returns whenever it is loads on a commodity with significant weight in the DJCI [21].

Furthermore, column (3) and (4) repeat the exercise but adding as control an open interest variable which captures additional omitted effects of the increasing capital inflows that are not perfectly captured by the coefficients Γ. Similarly, column (5) and (6) add as control a dollar open interest variable, i.e. open interest multiplied by the spot price. Overall, we see that the results are robust to the inclusion of these additional controls.

Taking together, the results here show that the index flow mechanism is a fundamental channel through which financialization has affected commodity futures strategies.

6 Conclusions

In this paper, we explore how the financialization and the influx of indexing has affected average returns to commodity futures investment strategies. We find that of the thirteen commodity futures strategies that had a significant average return before the financialization of this market, only two remain profitable post-financialization. We find that this decay in strategy returns is primarily the result of a dramatic fall in the price of risk to the systematic factors in this asset class.

Our results provide strong evidence in favour of the view that the financialization of commodity markets has affected certain commodity market participants. Our results are consistent with the model of Sockin and Xiong (2015) where information frictions impede commodity futures market participants from using prices and volume information in making profitable investment decisions.

[21] About 71bps p.m. if the strategies trade on the top-3 commodities for weights in the index (column 2 of Table 6), 162bps if the strategies trade on the commodity with the highest (i.e. top-1) weight in the DJCI (column 1).

References

Amihud, Yakov, Haim Mendelson, and Beni Lauterbach, 1997, Market microstructure and securities values: Evidence from the tel aviv stock exchange, *Journal of financial Economics* 45, 365–390.

Andrews, Donald WK, 1991, Heteroskedasticity and autocorrelation consistent covariance matrix estimation, *Econometrica: Journal of the Econometric Society* 817–858.

Asness, Clifford S, Tobias J Moskowitz, and Lasse Heje Pedersen, 2013, Value and momentum everywhere, *The Journal of Finance* 68, 929–985.

Baba Yara, Fahiz, Martijn Boons, and Andrea Tamoni, 2021, Value return predictability across asset classes and commonalities in risk premia, *Review of Finance* 25, 449–484.

Baker, Steven D, 2021, The financialization of storable commodities, *Management Science* 67, 471–499.

Bakshi, Gurdip, Xiaohui Gao, and Alberto G Rossi, 2019, Understanding the sources of risk underlying the cross section of commodity returns, *Management Science* 65, 619–641.

Bartram, Söhnke M, Leslie Djuranovik, and Anthony Garratt, 2022, Currency anomalies, *Available at SSRN 3194107* .

Basak, Suleyman, and Anna Pavlova, 2016, A model of financialization of commodities, *The Journal of Finance* 71, 1511–1556.

Basu, Devraj, and Joëlle Miffre, 2013, Capturing the risk premium of commodity futures: The role of hedging pressure, *Journal of Banking & Finance* 37, 2652–2664.

Ben-David, Itzhak, Francesco Franzoni, and Rabih Moussawi, 2018, Do etfs increase volatility?, *The Journal of Finance* 73, 2471–2535.

Bessembinder, Hendrik, 1992, Systematic risk, hedging pressure, and risk premiums in futures markets, *The Review of Financial Studies* 5, 637–667.

Bianchi, Robert J, Michael E Drew, and John Hua Fan, 2015, Combining momentum with reversal in commodity futures, *Journal of Banking & Finance* 59, 423–444.

Boons, Martijn, Frans De Roon, and Marta Szymanowska, 2014, The price of commodity risk in stock and futures markets, in *Afa 2012 Chicago meetings paper*.

Boons, Martijn, and Melissa Porras Prado, 2019, Basis-momentum, *The Journal of Finance* 74, 239–279.

Brennan, Michael J, 1958, The supply of storage, *The American Economic Review* 48, 50–72.

Brogaard, Jonathan, Matthew C Ringgenberg, and David Sovich, 2019, The economic impact of index investing, *The Review of Financial Studies* 32, 3461–3499.

Büyükşahin, Bahattin, and Michel A Robe, 2014, Speculators, commodities and cross-market linkages, *Journal of International Money and Finance* 42, 38–70.

CFTC, Commodity Futures Trading Commission, 2008, Staff report on commodity swap dealers index traders with commission recommendations .

Cheng, Ing-Haw, Andrei Kirilenko, and Wei Xiong, 2015, Convective risk flows in commodity futures markets, *Review of Finance* 19, 1733–1781.

Cheng, Ing-Haw, and Wei Xiong, 2014, Financialization of commodity markets, *Annu. Rev. Financ. Econ.* 6, 419–441.

Christoffersen, Peter, and Xuhui Nick Pan, 2018, Oil volatility risk and expected stock returns, *Journal of Banking & Finance* 95, 5–26.

Coles, Jeffrey L, Davidson Heath, and Matthew C Ringgenberg, 2022, On index investing, *Journal of Financial Economics* 145, 665–683.

Daniel, Kent, David Hirshleifer, and Avanidhar Subrahmanyam, 1998, Investor psychology and security market under-and overreactions, *the Journal of Finance* 53, 1839–1885.

De Bondt, Werner FM, and Richard Thaler, 1985, Does the stock market overreact?, *The Journal of finance* 40, 793–805.

De Roon, Frans A, Theo E Nijman, and Chris Veld, 2000, Hedging pressure effects in futures markets, *The Journal of Finance* 55, 1437–1456.

Deaton, Angus, and Guy Laroque, 1996, Competitive storage and commodity price dynamics, *Journal of Political Economy* 104, 896–923.

Dhume, Deepa, 2010, Using durable consumption risk to explain commodities returns, *Work. Pap., Harv. Univ* .

Erb, Claude B, and Campbell R Harvey, 2006, The strategic and tactical value of commodity futures, *Financial Analysts Journal* 62, 69–97.

Fama, Eugene F, and James D MacBeth, 1973, Risk, return, and equilibrium: Empirical tests, *Journal of political economy* 81, 607–636.

Fang, Xiang, Yang Liu, and Nikolai Roussanov, 2022, Getting to the core: Inflation risks within and across asset classes, Technical report, National Bureau of Economic Research.

Fernandez-Perez, Adrian, Bart Frijns, Ana-Maria Fuertes, and Joelle Miffre, 2018, The skewness of commodity futures returns, *Journal of Banking & Finance* 86, 143–158.

Gerakos, Joseph, and Juhani T Linnainmaa, 2018, Decomposing value, *The Review of Financial Studies* 31, 1825–1854.

Goldstein, Itay, and Liyan Yang, 2021, Commodity financialization and information transmission, *The Journal of Finance* forthcoming.

Gorton, Gary, and K Geert Rouwenhorst, 2006, Facts and fantasies about commodity futures, *Financial Analysts Journal* 62, 47–68.

Gorton, Gary B, Fumio Hayashi, and K Geert Rouwenhorst, 2013, The fundamentals of commodity futures returns, *Review of Finance* 17, 35–105.

Hamilton, James D, and Jing Cynthia Wu, 2014, Risk premia in crude oil futures prices, *Journal of International Money and Finance* 42, 9–37.

Hamilton, James D, and Jing Cynthia Wu, 2015, Effects of index-fund investing on commodity futures prices, *International economic review* 56, 187–205.

Henderson, Brian J, Neil D Pearson, and Li Wang, 2015, New evidence on the financialization of commodity markets, *The Review of Financial Studies* 28, 1285–1311.

Hicks, John R, 1939, Value and capital, *Cambridge, UK: Oxford Univ. Press* .

Hong, Harrison, and Motohiro Yogo, 2012, What does futures market interest tell us about the macroeconomy and asset prices?, *Journal of Financial Economics* 105, 473–490.

Hou, Kewei, Chen Xue, and Lu Zhang, 2020, Replicating anomalies, *The Review of Financial Studies* 33, 2019–2133.

Irwin, Scott H, and Dwight R Sanders, 2011, Index funds, financialization, and commodity futures markets, *Applied Economic Perspectives and Policy* 33, 1–31.

Israeli, Doron, Charles Lee, and Suhas A Sridharan, 2017, Is there a dark side to exchange traded funds? an information perspective, *Review of Accounting Studies* 22, 1048–1083.

Jensen, Theis Ingerslev, Bryan T Kelly, and Lasse Heje Pedersen, 2021, Is there a replication crisis in finance?, Technical report, National Bureau of Economic Research.

Kaldor, Nicholas, 1939, Speculation and economic stability, *The Review of Economic Studies* 7, 1–27.

Kang, Wenjin, K Geert Rouwenhorst, and Ke Tang, 2020, A tale of two premiums: the role of hedgers and speculators in commodity futures markets, *The Journal of Finance* 75, 377–417.

Keynes, John Maynard, 1930, Treatise on money: Pure theory of money vol. i.

Koijen, Ralph SJ, Tobias J Moskowitz, Lasse Heje Pedersen, and Evert B Vrugt, 2018, Carry, *Journal of Financial Economics* 127, 197–225.

Kozak, Serhiy, Stefan Nagel, and Shrihari Santosh, 2018, Interpreting factor models, *The Journal of Finance* 73, 1183–1223.

Lettau, Martin, and Markus Pelger, 2020, Factors that fit the time series and cross-section of stock returns, *The Review of Financial Studies* 33, 2274–2325.

Marshall, Ben R, Nhut H Nguyen, and Nuttawat Visaltanachoti, 2012, Commodity liquidity measurement and transaction costs, *The Review of Financial Studies* 25, 599–638.

Marshall, Ben R, Nhut H Nguyen, and Nuttawat Visaltanachoti, 2013, Liquidity commonality in commodities, *Journal of Banking & Finance* 37, 11–20.

McCracken, Michael W, and Serena Ng, 2016, Fred-md: A monthly database for macroeconomic research, *Journal of Business & Economic Statistics* 34, 574–589.

McLean, R David, and Jeffrey Pontiff, 2016, Does academic research destroy stock return predictability?, *The Journal of Finance* 71, 5–32.

Melone, Alessandro, Otto Randl, Leopold Sögner, and Josef Zechner, 2021, Stock-oil comovement: Fundamentals or financialization?, *Available at SSRN 3668239* .

Miffre, Joëlle, and Georgios Rallis, 2007, Momentum strategies in commodity futures markets, *Journal of Banking & Finance* 31, 1863–1886.

Newey, Whitney K, and Kenneth D West, 1987, A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica: Journal of the Econometric Society* 703–708.

Shanken, Jay, 1992, On the estimation of beta-pricing models, *The review of financial studies* 5, 1–33.

Singleton, Kenneth J, 2014, Investor flows and the 2008 boom/bust in oil prices, *Management Science* 60, 300–318.

Sockin, Michael, and Wei Xiong, 2015, Informational frictions and commodity markets, *The Journal of Finance* 70, 2063–2098.

Stoll, Hans R, and Robert E Whaley, 2010, Commodity index investing and commodity futures prices, *Journal of Applied Finance (Formerly Financial Practice and Education)* 20.

Szymanowska, Marta, Frans De Roon, Theo Nijman, and Rob Van Den Goorbergh, 2014, An anatomy of commodity futures risk premia, *The Journal of Finance* 69, 453–482.

Working, Holbrook, 1949, The theory of price of storage, *The American Economic Review* 39, 1254–1262.

7 Tables and Figures

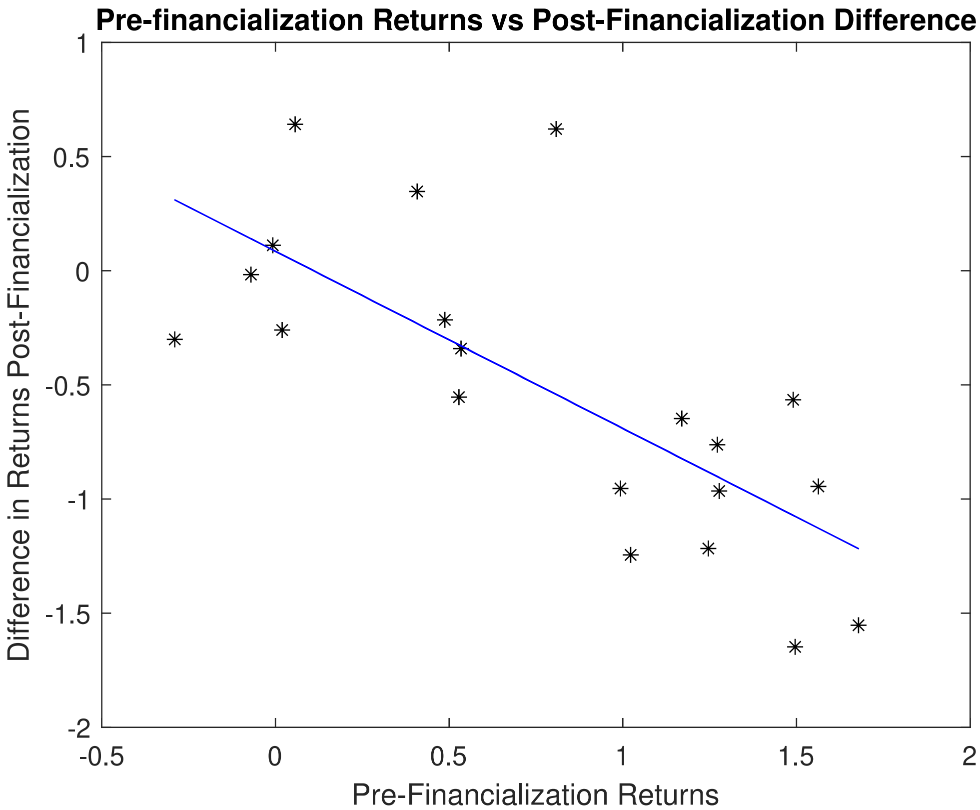

Figure 1: Relation between pre- and post-financializtion returns

This scatter plot shows the relation between the monthly returns to the 20 commodity trading strategies pre-financialization and the changes in their returns post-financialization. The returns pre-financialization are mean monthly excess returns in percentage points (i.e % per month). Changes in returns post-financialization are instead the difference of the mean monthly excess returns in percentage points between post-financialization and pre-financialization returns. The sample is monthly from March 1986 to August 2021. The commodity investment strategies are described in Section A.2 of the Appendix and their performance across the different subsamples (Full Sample, Pre- and Post-Financialization periods) are analyzed in Table 1.

Table 1: Descriptive Statistics - Returns to the Commodity Investment Strategies (Over Subsamples)

This table reports the returns to the commodity investment strategies built on the characteristics described in Section 1.2. The construction of the excess returns takes into account the first notice day convention following Bakshi et al. (2019). We build end-of-month series for commodity returns from March 1986 to August 2021. These data are collected from CRB, Datastream, and Factset. The left panel reports the statistics for the returns to the strategies over the full sample of data (03/1986 to 08/2021). While, the middle and right panels report the descriptive statistics for the strategies, respectively, over the pre- and post-financialization periods, where the sample is split around January 2004. The choice of 2004 as date for the financialization of commodity markets is driven by the previous literature (see Boons et al. (2014) and Basak and Pavlova (2016), among others). The data for the sorting variables are retrieved from different sources, as reported in the separate Appendix A.2 which also describes how the strategies are constructed. *Inventory* is available only up to beginning of 2011. Average returns (*Mean*) and standard deviations (*Std*) are annualized in percentage. *SR* refers to sharpe ratio. We compute test statistics (*tstat*) using Newey and West (1987) corrected standard errors (with lag selection following Andrews (1991)). We denote with ***, **, * estimates significant at the, respectively, 1%, 5% and 10% level.

Full Sample					Pre-Financialization					Post-Financialization				
Factor	Mean%	Std%	SR	tstat	Factor	Mean%	Std%	SR	tstat	Factor	Mean%	Std%	SR	tstat
Carry***	14.82	18.76	0.79	4.81	Carry***	18.49	20.64	0.90	3.95	Carry***	11.10	16.62	0.67	2.69
Skewn***	13.86	20.83	0.67	4.00	Skewn*	10.63	22.89	0.46	1.89	Skewn***	17.14	18.52	0.93	4.13
BaMom***	13.13	25.29	0.52	2.84	BaMom***	18.77	25.77	0.73	3.17	BaMom	7.42	24.74	0.30	1.08
Inven**	-7.19	13.93	-0.52	-2.55	Inven**	-8.30	14.26	-0.58	-2.36	Inven	-4.37	13.07	-0.33	-0.94
CVDhu**	7.51	17.34	0.43	2.54	CVDhu***	14.56	17.69	0.82	3.76	CVDhu	0.35	16.77	0.02	0.09
Mom06**	9.14	23.97	0.38	2.48	Mom06**	11.98	25.93	0.46	2.05	Mom06	6.27	21.83	0.29	1.40
Mom01**	10.23	25.05	0.41	2.39	Mom01**	14.28	26.28	0.54	2.34	Mom01	6.12	23.74	0.26	1.11
Volat**	8.35	23.43	0.36	2.34	Volat***	18.33	25.03	0.73	3.89	Volat	-1.82	21.34	-0.09	-0.36
MoB12**	10.24	28.67	0.36	2.16	MoB12***	18.85	31.26	0.60	2.63	MoB12	1.51	25.61	0.06	0.27
Mom03**	8.83	24.41	0.36	2.08	Mom03**	13.84	25.27	0.55	2.25	Mom03	3.75	23.48	0.16	0.69
InflB*	7.07	21.80	0.32	1.85	InflB	5.10	19.98	0.26	1.09	InflB	9.06	23.54	0.38	1.54
Averg*	4.53	12.00	0.38	1.81	Averg***	5.77	9.47	0.61	2.58	Averg	3.27	14.12	0.23	0.77
Rever*	6.07	22.32	0.27	1.68	Rever**	11.58	23.02	0.50	2.25	Rever	0.47	21.51	0.02	0.09
DollB	-4.67	19.71	-0.24	-1.41	DollB	-2.29	18.81	-0.12	-0.50	DollB	-7.09	20.60	-0.34	-1.48
Mom12	4.83	21.37	0.23	1.38	Mom12**	12.22	22.92	0.53	2.40	Mom12	-2.67	19.50	-0.14	-0.61
HedPr	3.83	18.17	0.21	1.26	HedPr	-0.66	18.24	-0.04	-0.15	HedPr**	8.38	18.04	0.46	2.15
CVSzy	3.78	17.55	0.22	1.18	CVSzy	5.23	19.33	0.27	1.08	CVSzy	2.32	15.57	0.15	0.58
Basis	2.57	25.74	0.10	0.58	Basis	5.40	27.94	0.19	0.78	Basis	-0.31	23.33	-0.01	-0.06
OpeIn	-1.33	18.20	-0.07	-0.43	OpeIn	0.19	21.03	0.01	0.04	OpeIn	-2.89	14.83	-0.19	-0.87
Liqui	-0.60	15.57	-0.04	-0.20	Liqui	-0.15	16.02	-0.01	-0.04	Liqui	-1.05	15.14	-0.07	-0.27
Value	0.59	18.14	0.03	0.18	Value	-0.05	18.40	0.00	-0.01	Value	1.24	17.91	0.07	0.29

Table 2: Unconditional Asset Pricing Tests (Over Subsamples) - Second Stage Regressions

This table reports the results for the second (cross-sectional) stage of the Fama and MacBeth (1973) asset pricing tests. We use as test assets the returns to the commodity investment strategies presented in Table 1; while, the six candidate factors are the six RP-PCs extracted as in Lettau and Pelger (2020). *Panel A* reports the results for the test conducted over the full sample of data (03/1986 to 08/2021); while, *Panel B* and *Panel C* report the results for the tests conducted, respectively, over the pre- and post-financialization periods, where the sample is split around January 2004. The choice of 2004 as date for the financialization of commodity markets is driven by the previous literature (see Boons et al. (2014) and Basak and Pavlova (2016), among others). Mean (*Mean*) and prices of risk (*RP*) for each latent factor (the RP-PCs), as well as for the estimated intercepts, are reported in annualized percentage points. The test-statistics are computed using Newey and West (1987)- ($tstat_{nw}$) and Shanken (1992)-corrected ($tstat_{sh}$) standard errors. Cross-sectional R^2 are in percentage points. The risk premium parameter of the Lettau and Pelger (2020) procedure is set equal to 10. The sample is monthly from March 1986 to August 2021. Results for the first (time-series) stage of the asset pricing tests are reported in Table A.3 in the Appendix.

Panel A	Full Sample						
Factors	Intercept	PC1	PC2	PC3	PC4	PC5	PC6
Mean (%)		30.77	15.71	5.76	0.39	4.26	4.37
RP (%)	0.84	27.94	14.09	4.38	0.92	3.47	4.79
$tstat_{nw}$	[0.77]	[3.41]	[3.31]	[0.78]	[0.22]	[0.83]	[1.09]
$tstat_{sh}$	[0.81]	[3.18]	[2.74]	[0.86]	[0.20]	[0.82]	[1.19]
R2 (%)		33.69	78.58	84.26	84.22	87.30	90.47

Panel B	Pre-Financialization						
Factors	Intercept	PC1	PC2	PC3	PC4	PC5	PC6
Mean (%)		46.09	15.67	14.51	0.99	2.25	-5.55
RP (%)	0.45	44.04	13.40	13.69	1.55	0.97	-6.07
$tstat_{nw}$	[0.45]	[3.80]	[2.70]	[1.76]	[0.36]	[0.14]	[-1.12]
$tstat_{sh}$	[0.31]	[3.24]	[1.79]	[1.87]	[0.25]	[0.16]	[-1.03]
R2 (%)		40.24	59.24	80.28	80.40	80.56	84.75

Panel C	Post-Financialization						
Factors	Intercept	PC1	PC2	PC3	PC4	PC5	PC6
Mean (%)		15.88	15.30	-3.29	-0.42	5.96	14.26
RP (%)	1.39	11.55	12.29	-5.31	-0.06	5.76	14.26
$tstat_{nw}$	[0.89]	[1.42]	[1.84]	[-0.80]	[-0.01]	[1.24]	[2.61]
$tstat_{sh}$	[1.06]	[1.07]	[1.73]	[-0.76]	[-0.01]	[0.98]	[2.60]
R2 (%)		6.31	39.63	41.28	41.30	47.89	77.64

Table 3: Latent Factors and Their Relations to the Commodity Investment Strategies

This table shows the regression results based on univariate regressions of each latent factor (i.e. the RP-PCs extracted as in Lettau and Pelger (2020)) on a constant and the returns to a commodity investment strategy (i.e. $RPPC_t^j = \psi_0 + \psi_1 r_t^{strategy-i} + u_t$). In the table, *Cons* refers to ψ_0, while *Slope* to ψ_1. The column *R2* reports the R^2 in percentage. While, the column *Corr* reports the correlation coefficient between the strategy and the RP-PC. Standard errors are computed with Newey and West (1987) adjustment (number of lags selected as in Andrews (1991)). We build end-of-month series for commodity returns from March 1986 to August 2021. We denote with ***, **, * estimates significant at the, respectively, 1%, 5% and 10% level. The construction of the strategies is described in the separate Appendix (A.2).

	PC1				PC2				PC3				PC4				PC5				PC6			
Variables	Cons	Slope	R2	Corr	Cons	Slope	R2	Corr	Cons	Slope	R2	Corr	Cons	Slope	R2	Corr	Cons	Slope	R2	Corr	Cons	Slope	R2	Corr
Averg	0.02***	0.87***	4.52	0.21	0.01***	1.03***	17.41	0.42	0.01	-0.46**	3.62	-0.19	0.00	-0.58***	6.88	-0.26	0.00	-0.22	1.12	-0.11	0.01*	-0.62***	10.19	-0.32
Carry	0.01	1.41***	29.32	0.54	0.01**	0.27	2.86	0.17	0.01**	-0.36	5.33	-0.23	0.00	0.27***	3.74	0.19	0.00	0.00	0.00	0.00	0.00	0.03	0.08	0.03
Mom06	0.01***	1.62***	63.11	0.79	0.02***	-0.53***	18.53	-0.43	0.00	0.00	0.00	0.00	0.00	-0.16**	2.02	-0.14	0.00	-0.17**	2.79	-0.17	0.00	-0.04	0.13	-0.04
Mom12	0.02***	1.79***	60.82	0.78	0.01***	-0.36***	6.77	-0.26	0.01	-0.10	0.49	-0.07	0.00	0.20*	2.57	0.16	0.01*	-0.55***	22.42	-0.47	0.00	-0.27***	6.07	-0.25
InflB	0.02***	0.52***	5.36	0.23	0.01***	0.51***	14.16	0.38	0.01***	-0.70***	27.34	-0.52	0.00	-0.25**	4.10	-0.20	0.00	-0.02	0.03	-0.02	0.00	-0.22**	4.11	-0.20
DollB	0.03***	-0.14	0.32	-0.06	0.01***	-0.48***	10.40	-0.32	0.01**	0.78***	27.55	0.52	0.00	0.14	1.13	0.11	0.00	-0.30***	5.61	-0.24	0.00	0.02	0.02	0.01
CVSzy	0.02***	1.01***	13.17	0.36	0.02***	-0.67***	15.82	-0.40	0.01	-0.16	0.95	-0.10	0.00	-0.25**	2.81	-0.17	0.00	0.37***	6.94	0.26	0.00	0.01	0.00	0.01
CVDhu	0.02***	0.97***	11.80	0.34	0.01***	0.14	0.72	0.08	0.00	0.02	0.02	0.01	0.00	-0.23***	2.30	-0.15	0.01*	-0.37***	6.64	-0.26	0.00	0.00	0.00	0.00
HedPr	0.02***	0.42*	2.41	0.16	0.01***	0.06	0.12	0.03	0.01	-0.24**	2.19	-0.15	0.00	-0.25***	3.02	-0.17	0.00	-0.27***	3.86	-0.20	0.00	0.60***	21.50	0.46
BaMom	0.02***	0.95***	23.89	0.49	0.01***	0.02	0.03	0.02	0.01	-0.19**	2.61	-0.16	-0.01***	0.69***	43.15	0.66	0.00	0.31***	9.78	0.31	0.01	-0.14	2.20	-0.15
MoB12	0.01***	1.46***	73.03	0.85	0.02***	-0.43***	17.33	-0.42	0.00	0.04	0.18	0.04	0.00	0.12	1.66	0.13	0.01*	-0.29***	11.01	-0.33	0.01*	-0.19**	5.41	-0.23
Basis	0.03***	-1.01***	27.98	-0.53	0.01***	0.44***	14.77	0.38	0.00	0.67***	34.48	0.59	0.00	-0.25***	5.86	-0.24	0.00	0.03	0.08	0.03	0.00	-0.14*	2.22	-0.15
Volat	0.02***	0.51**	5.98	0.24	0.01***	0.57***	20.12	0.45	0.00	0.15	1.47	0.12	0.00	-0.35***	9.32	-0.31	0.01*	-0.31***	8.45	-0.29	0.01**	-0.46***	21.32	-0.46
Value	0.03***	-1.43***	28.00	-0.53	0.01***	0.42***	6.80	0.26	0.00	0.59***	13.62	0.37	0.00	0.24**	2.61	0.16	0.00	0.47***	11.76	0.34	0.00	-0.04	0.08	-0.03
Skewn	0.02***	0.30	1.65	0.13	0.00	0.83***	34.12	0.58	0.01	-0.06	0.16	-0.04	0.00	-0.03	0.05	-0.02	0.01*	-0.29***	5.98	-0.24	0.00	0.41***	13.13	0.36
OpeIn	0.03***	0.41	2.28	0.15	0.01***	-0.01	0.01	-0.01	0.00	-0.30***	3.49	-0.19	0.00	-0.32***	4.84	-0.22	0.00	0.53***	15.04	0.39	0.00	-0.79***	37.39	-0.61
Liqui	0.03***	0.41	1.74	0.13	0.01***	0.10	0.30	0.05	0.00	-0.65***	12.10	-0.35	0.00	-0.32**	3.59	-0.19	0.00	0.63***	15.93	0.40	0.00	-0.68***	20.69	-0.45
Mom03	0.01***	1.53***	58.28	0.76	0.02***	-0.56***	21.38	-0.46	0.00	-0.01	0.00	0.00	0.00	-0.36***	11.11	-0.33	0.00	0.18*	3.12	0.18	0.00	0.03	0.09	0.03
Mom01	0.02***	1.12***	32.90	0.57	0.02***	-0.53***	20.52	-0.45	0.00	0.23***	3.91	0.20	0.00	-0.41***	14.57	-0.38	0.00	0.33***	10.96	0.33	0.00	0.14	2.39	0.15
Rever	0.02***	0.29*	1.76	0.13	0.01***	-0.14	1.09	-0.10	0.00	0.93***	50.60	0.71	0.00	0.35***	8.40	0.29	0.00	0.12	1.18	0.11	0.00	-0.17**	2.69	-0.16

Table 4: Latent Factors and Their Relations to Macro-Financial Variables

This table shows the regression results based on univariate regressions of each latent factor (i.e. the RP-PCs extracted as in Lettau and Pelger (2020)) on a constant and a macro-financial variable (i.e. $RPPC_t^j = \phi_0 + \phi_1 \Delta MacroFin_t^i + \epsilon_t$). In the table, *Cons* refers to ϕ_0, while *Slope* to ϕ_1. The column *R2* reports the R^2 in percentage. We build end-of-month series for commodity returns from March 1986 to August 2021. Standard errors are computed with Newey and West (1987) adjustment (number of lags selected as in Andrews (1991)). We denote with ***, **, * estimates significant at the, respectively, 1%, 5% and 10% level. Table A.5 in the separate Appendix reports the description of each variable.

		PC1			PC2			PC3			PC4			PC5			PC6		
Variables	Category	Cons	Slope	R2	Cons	Slope	R2	Cons	Slope	R2	Cons	Slope	R2	Cons	Slope	R2	Cons	Slope	R2
sliq	Financial	0.03***	-0.05**	1.57%	0.01***	-0.02	0.51%	0.00	0.01	0.26%	0.00	0.01	0.40%	0.00	0.01	0.13%	0.00	0.00	0.00%
icap	Financial	0.03***	-0.14	0.47%	0.01***	0.22***	3.11%	0.00	-0.13	1.10%	0.00	0.07	0.36%	0.00	-0.06	0.33%	0.00	-0.04	0.15%
gfc	Financial	0.03***	0.05	0.61%	0.01***	0.08***	5.45%	0.00	-0.06***	3.13%	0.00	0.00	0.02%	0.00	-0.03	0.69%	0.00	-0.02	0.53%
ted	Financial	0.03***	0.03	0.20%	0.01***	-0.04***	1.09%	0.00	0.04**	1.35%	0.00	0.00	0.01%	0.00	0.04**	1.58%	0.00	-0.03	0.76%
GS10	Financial	0.03***	-0.01	0.06%	0.01***	0.05***	1.48%	0.00	-0.03	0.66%	0.00	-0.01	0.14%	0.00	0.04**	1.28%	0.00	-0.02	0.52%
TWEXAFE	Financial	0.03***	-0.91**	1.13%	0.01***	-0.39	0.58%	0.00	0.72***	1.99%	0.00	0.22	0.22%	0.00	-0.18	0.18%	0.00	0.42*	1.04%
BAAMAAA	Financial	0.03***	-0.05	0.15%	0.01***	-0.13*	2.42%	0.00	0.14***	3.13%	0.00	0.05	0.42%	0.00	0.01	0.02%	0.00	0.05	0.67%
S&P 500	Financial	0.03***	0.19	0.25%	0.01***	0.37***	2.60%	0.01	-0.29*	1.63%	0.00	0.07	0.12%	0.00	-0.29***	2.31%	0.00	-0.09	0.23%
equ_vol	Financial	0.02***	-5.66**	2.48%	0.01***	-0.94	0.19%	0.01	3.61***	2.82%	0.00	0.05	0.00%	0.00	2.35***	1.67%	0.00	0.11	0.00%
comm_vol	Financial	0.02***	0.37	0.00%	0.01***	-5.59*	1.58%	0.01	9.20**	4.43%	0.00	-1.10	0.08%	0.00	4.29*	1.36%	0.00	2.12	0.36%
gecon	Macro	0.03***	0.00	0.00%	0.01***	0.06***	6.51%	0.00	-0.04**	2.23%	0.00	-0.01	0.22%	0.00	-0.03***	2.24%	0.00	-0.02	1.05%
INDPRO	Macro	0.03***	0.02	0.00%	0.01***	1.00	0.77%	0.00	-0.04	0.00%	0.00	-0.32	0.10%	0.00	-1.17***	1.52%	0.01	-0.97**	1.15%
WPSFD49207	Macro	0.03***	1.28	0.46%	0.01***	1.89*	2.81%	0.00	-2.12***	3.62%	0.00	-0.81	0.64%	0.00	-0.69	0.54%	0.00	-0.60	0.45%
CPIAUCSL	Macro	0.03***	4.84*	0.85%	0.01***	4.79**	2.31%	0.00	-6.74***	4.67%	0.00	-4.08***	2.08%	0.00	-2.70**	1.05%	0.00	-2.08	0.69%

Table 5: Regression of Factors on Post-Financialization and Post-Publication Indicators

This table reports the results from regressions of the returns to the commodity strategies (in percentage per month) on a dummy variable for the post-financialization period, a dummy variable for the period between financialization and the publication of the factor, and a dummy variable for the post-publication period. *Post-Financialization* is equal to one if the month is after the financialization of commodity markets (i.e. post 01/2004) and zero otherwise. *Post-FinaToPublication* is equal to one if the month is after the financialization of commodity markets (i.e. post 01/2004) but before the official publication date, and zero otherwise. *Post-Publication* is equal to one if the month is after the official publication date and zero otherwise. The choice of 2004 as date for the financialization of commodity markets is driven by the previous literature (see Boons et al. (2014) and Basak and Pavlova (2016), among others). The data contain monthly series from March 1986 to August 2021. Regressions include factor fixed effects as indicated in the table. Standard errors are clustered on time. We denote with ***, **, * estimates significant at the, respectively, 1%, 5% and 10% level. The mean factor return pre-financialization is 0.784 (i.e. 78.4 bps per month).

	Factors			
	(1)	(2)	(3)	(4)
Post-Financialization	-0.511**		-0.438**	
Post-FinaToPublication				-0.437**
Post-Publication		-0.460**	-0.144	-0.577**
Observations	8,499	8,499	8,499	8,499
Factor Fixed Effect	Yes	Yes	Yes	Yes

Table 6: Index Flow Mechanism

This table reports the results from regressions of the returns to the commodity strategies (in percentage per month) on a dummy variable that takes value one if the futures strategy-i at time-t has exposure to any commodity in the top-n weighted commodities in the index at time t, and a dummy that at t is equal to one for all strategies that have at least a commodity with top-n exposure to the index. All regressions include factor fixed effect. Column (1) and (2) reports results when we restrict the commodities to have, respectively, top-1 and top-3 weights in the DJCI Index over the (monthly) period 01/2004 to 08/2021 (i.e. the post-financialization period). Column (3) and (4) repeat the same excercise but adding open interest as control variable; while, Column (5) and (6) add dollar open interest as control variable. The choice of 2004 as date for the financialization of commodity markets is driven by the previous literature (see Boons et al. (2014) and Basak and Pavlova (2016), among others). Standard errors are clustered on time. Cross-sectional R^2 are in percentage points. We denote with ***, **, * estimates significant at the, respectively, 1%, 5% and 10% level.

	Factors					
	(1)	(2)	(3)	(4)	(5)	(6)
$D_{i,t}$	-1.623***	-0.707***	-1.335***	-0.568*	-1.299***	-0.549**
Observations	4139	4139	4139	4139	3933	3933
Factor Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Exposure by Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Control for Open Interest	No	No	Yes	Yes	No	No
Control for Dollar Open Interest	No	No	No	No	Yes	Yes

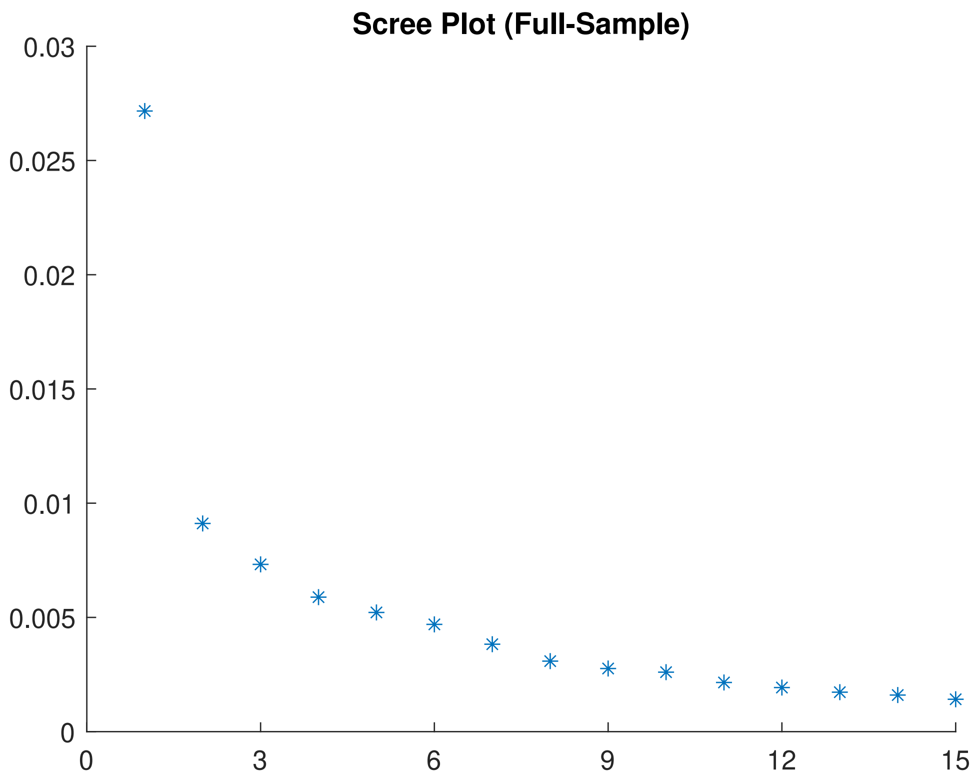

Figure 2: First 15 eigenvalues

This figure plots the first 15 eigenvalues of the data, arising from the application of the RP-PCA methodology (Lettau and Pelger (2020)) to the returns to the commodity investment strategies. The data cover the period 03/1986-08/2021.

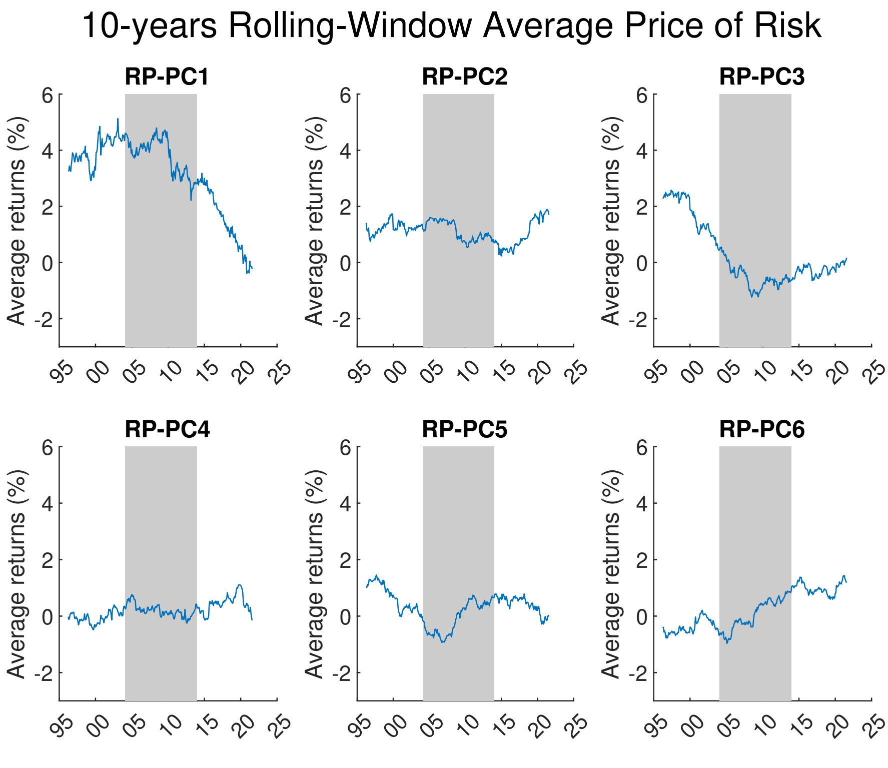

Figure 3: Rolling-window average price of risk
This figure plots the ten-year rolling average return of each PC in our six factor model. The grey rectangle starts at January 2004 (the financialization date) and ends in December 2013. The choice of 2004 as date for the financialization of commodity markets is driven by the previous literature (see Boons et al. (2014) and Basak and Pavlova (2016), among others). The data contain monthly series from March 1986 to August 2021.

A Appendix

A.1 Descriptive Statistics - Individual Commodities

Table A.1: **Descriptive Statistics - Excess Returns to Individual Commodities**
This table reports the descriptive statistics of the individual commodity futures excess returns computed as in equation (1). For each commodity, we report the annualized average returns (*Mean*), annualized standard deviation (*Std*), annualized Sharpe Ratios (*SR*) and the skewness (*Skew*) of the monthly returns, as well as the number of observations (N). The construction of the excess returns takes into account the first notice day convention following Bakshi et al. (2019). We build end-of-month series for commodity returns from March 1986 to August 2021. These data are collected from CRB, Datastream, and Factset.

	Mean %	Std %	SR	Skew	N
Crude oil	10.50	35.88	0.29	0.45	425
Gasoline	22.32	39.63	0.56	0.28	425
Heating oil	13.96	35.98	0.39	0.85	425
Natural gas	-8.04	47.14	-0.17	0.60	375
Gas-oil petroleum	10.87	33.09	0.33	0.21	383
Propane	27.20	64.15	0.42	7.30	263
Rough rice	-4.20	25.82	-0.16	1.02	416
Sugar	6.14	30.70	0.20	0.30	425
Corn	-2.96	25.91	-0.11	0.72	425
Oats	2.47	32.80	0.08	2.37	425
Wheat	-2.56	26.33	-0.10	0.42	425
Canola	1.72	20.79	0.08	0.08	422
Barley	0.87	20.12	0.04	0.29	278
Cotton	2.34	25.70	0.09	0.24	425
Lumber	0.42	32.87	0.01	0.70	425
Rubber	3.46	36.46	0.09	0.44	354
Feeder cattle	3.80	14.28	0.27	-0.13	425
Live cattle	2.48	13.69	0.18	-0.44	425
Lean hogs	-0.35	25.26	-0.01	-0.27	425
Pork bellies	4.06	37.65	0.11	0.56	304
Gold	2.73	15.30	0.18	0.17	425
Silver	4.40	27.92	0.16	0.38	425
Copper	12.07	25.79	0.47	0.17	425
Palladium	14.28	30.92	0.46	0.38	421
Platinum	4.83	21.89	0.22	-0.03	425
Soybeans oil	0.42	23.85	0.02	0.19	425
Soybeans meal	10.10	25.35	0.40	0.44	425
Soybeans	4.69	22.78	0.21	-0.02	425
Coffee	-2.83	35.67	-0.08	1.09	425
Orange juice	0.52	29.67	0.02	0.52	425
Cocoa	-1.67	27.77	-0.06	0.44	425
Milk	5.70	28.87	0.20	1.05	303

A.2 Variables Construction

1. Level (*Averg*): We follow Bakshi et al. (2019) in constructing the level factor (i.e. the average factor) as the excess returns of a strategy that goes long in all the available commodity futures.

2. Carry (*Carry*): We follow Bakshi et al. (2019) in constructing carry by sorting on the log of the slope of the futures curve (i.e. $log(y_t)$, with $y_t = \frac{F_t^{(1)}}{F_t^{(0)}}$) and in allocating commodities into 4 portfolios. Hence, commodities are sorted from most in contango (highest $ln(y_t) > 0$) to most backwardated (lowest $ln(y_t) > 0$).

3. Basis (*Basis*): We follow Boons and Prado (2019) in constructing basis by sorting on $B_t = \frac{(F_t^{(2)} - F_t^{(1)})}{F_t^{(1)}}$ and allocating commodities into 3 portfolios. The High (respectively, Low) portfolio contains the four commodities with the highest (respectively, lowest) ranked signal, while the Medium portfolio contains all remaining commodities.

4. Momentum 1-months (*Mom01*): We follow Miffre and Rallis (2007) in constructing momentum by sorting on the returns over the previous one-month. Commodities are then allocated in 5 portfolios.

5. Momentum 3-months (*Mom03*): We follow Miffre and Rallis (2007) in constructing momentum by sorting on the returns over the previous three-months. Commodities are then allocated in 5 portfolios.

6. Momentum 6-months (*Mom06*): We follow Bakshi et al. (2019) in constructing momentum by sorting on the past six-month performance. Commodities are then allocated in 5 portfolios.

7. Momentum 12-months (*Mom12*): We follow Szymanowska et al. (2014) in constructing (long-term) momentum by sorting on the cumulative log return from month $t - 12$ to $t - 1$ and allocating commodities into 4 portfolios.

8. Momentum 12-months (*MoB12*): We follow Boons and Prado (2019) in constructing (long-term) momentum by sorting on the cumulative log return from month $t - 11$

to t and allocating commodities into 3 portfolios. The High (respectively, Low) portfolio contains the four commodities with the highest (respectively, lowest) ranked signal, while the Medium portfolio contains all remaining commodities.

9. Reversal (*Rever*): Bianchi et al. (2015) show that a consistent reversal pattern is pronounced from month 12 to 30. We construct the contrarian strategy on a signal based on portfolio formation months 36-13, and allocate commodities into 5 portfolios.

10. Basis-Momentum (*BaMom*): We follow Boons and Prado (2019) in constructing basis-momentum by sorting on:

$$BM_t = \prod_{j=t-11}^{t} (1 + R_j^{(1)}) - \prod_{j=t-11}^{t} (1 + R_j^{(2)})$$

i.e. on the momentum between two consecutive nearby futures strategies and allocating commodities into 3 portfolios. The High (respectively, Low) portfolio contains the four commodities with the highest (respectively, lowest) ranked signal, while the Medium portfolio contains all remaining commodities.

11. Coefficient of variation using spot prices (*CVDhu*): We follow Dhume (2010) in constructing the coefficient of variation as the variance of the past three months daily spot prices scaled by their mean. Commodities are then allocated into 5 portfolios using the demeaned values (where the mean is computed over the previous 60 months).

12. Coefficient of variation using returns (*CVSzy*): We follow Szymanowska et al. (2014) in constructing the coefficient of variation as the variance on the past daily returns scaled by the mean return and allocating commodities into 4 portfolios.

13. Volatility (*Volat*): We follow Gorton et al. (2013) in constructing volatility as the square root of the average squared daily excess returns of the month over which the excess return is calculated, multiplied by the square root of 365. Thus, this measure

is forward-looking. Moreover, volatility is demeaned at the commodity level. We allocate commodities into 4 portfolios.

14. Inventory (*Inven*): We refer the interested reader to Section 3.2 and Appendix B of Gorton et al. (2013) for how this variable is constructed. Our data end at the beginning of 2011. Following their paper, commodities are allocated into 2 portfolios.

15. Hedging pressure (*HedPr*): We follow Szymanowska et al. (2014) and Basu and Miffre (2013) in constructing hedging pressure (for hedgers) as the difference between the number of short and long hedge positions by large traders in proportion to the total number of hedge positions by large traders in that market:

$$hp_t = \frac{\#\text{short hedge positions} - \#\text{long hedge positions}}{\text{total } \# \text{ hedge positions}}$$

The positions are measured by the number of contracts in the market. The data are retrieved from the Commitment of Traders reports issued by the Commodity Futures Trading Commission (CFTC). Commodities are then allocated into 4 portfolios.

16. Open interest (*OpeIn*): We follow Hong and Yogo (2012) in constructing open interest as the total open interest in futures market. We allocate commodities into 4 portfolios. The data are retrieved from the Commitment of Traders reports issued by the Commodity Futures Trading Commission (CFTC).

17. Liquidity (*Liqui*): We follow Marshall et al. (2012) and Marshall et al. (2013) in constructing liquidity as the Amivest measure for liquidity of Amihud et al. (1997), i.e. as the volume on a trading day divided by the absolute value of the daily return. We allocate commodities into 4 portfolios.

18. Value (*Value*): We follow Asness et al. (2013) in constructing value as the log of the spot price 5 years ago (actually, of the average spot price from 4.5 to 5.5 years ago) divided by the most recent spot price and allocating commodities into 3 portfolios. Hence, value can be seen as the negative of the spot return over the last 5 years.

19. Inflation-β (*InflB*): we sort commodities based on the betas estimated from a 60-month rolling window regression of monthly commodity futures returns on changes in one-month CPI inflation. We then allocate commodities into 4 portfolios (see also Szymanowska et al. (2014)).

20. Dollar-β (*DollB*): we sort commodities based on the betas estimated from a 60-month rolling window regression of monthly commodity futures returns on changes in a broad US dollar index. We then allocate commodities into 4 portfolios (see also Szymanowska et al. (2014)).

21. Skewness (*Skewn*): We follow Fernandez-Perez et al. (2018) in constructing skewness by sorting on the coefficient of skewness of the daily commodity returns from month $t - 11$ to t. We allocate commodities into 5 portfolios.

A.3 Average Returns to the Commodity Trading Strategies (Across Subsamples)

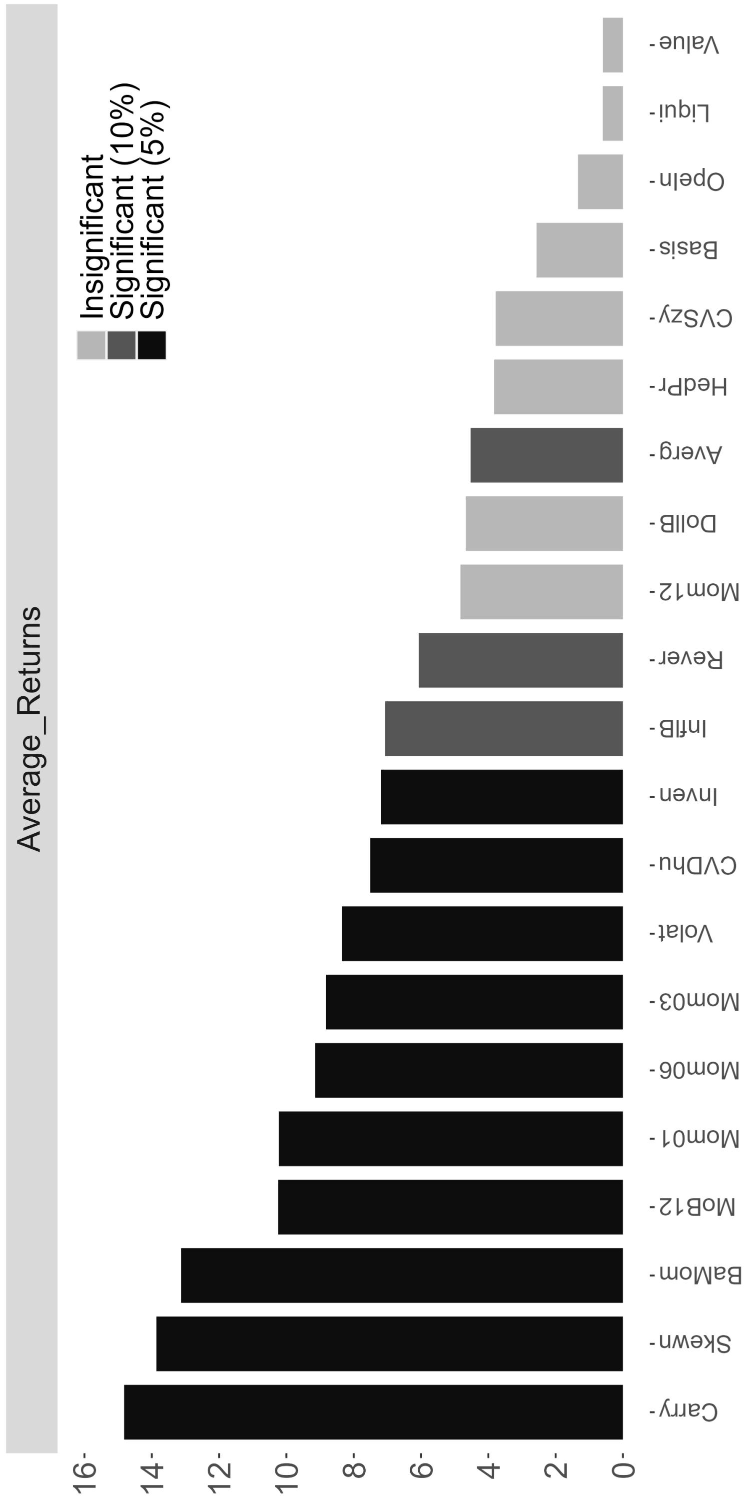

Figure A.1: Average returns to the commodity investment strategies (Full Sample)
This histogram shows the average returns to each of the commodity trading strategies. The returns are annualized excess returns in percentage points. The sample is monthly from March 1986 to August 2021. Black and dark grey bars represent, respectively, strategies that deliver returns significant at the 5% and 10% significant level; while, light grey bars represent strategies that deliver average returns not statistically significant. The commodity investment strategies are described in Section A.2 of the Appendix and their performance across the different subsamples (Full Sample, Pre- and Post-Financialization periods) are analyzed in Table 1.

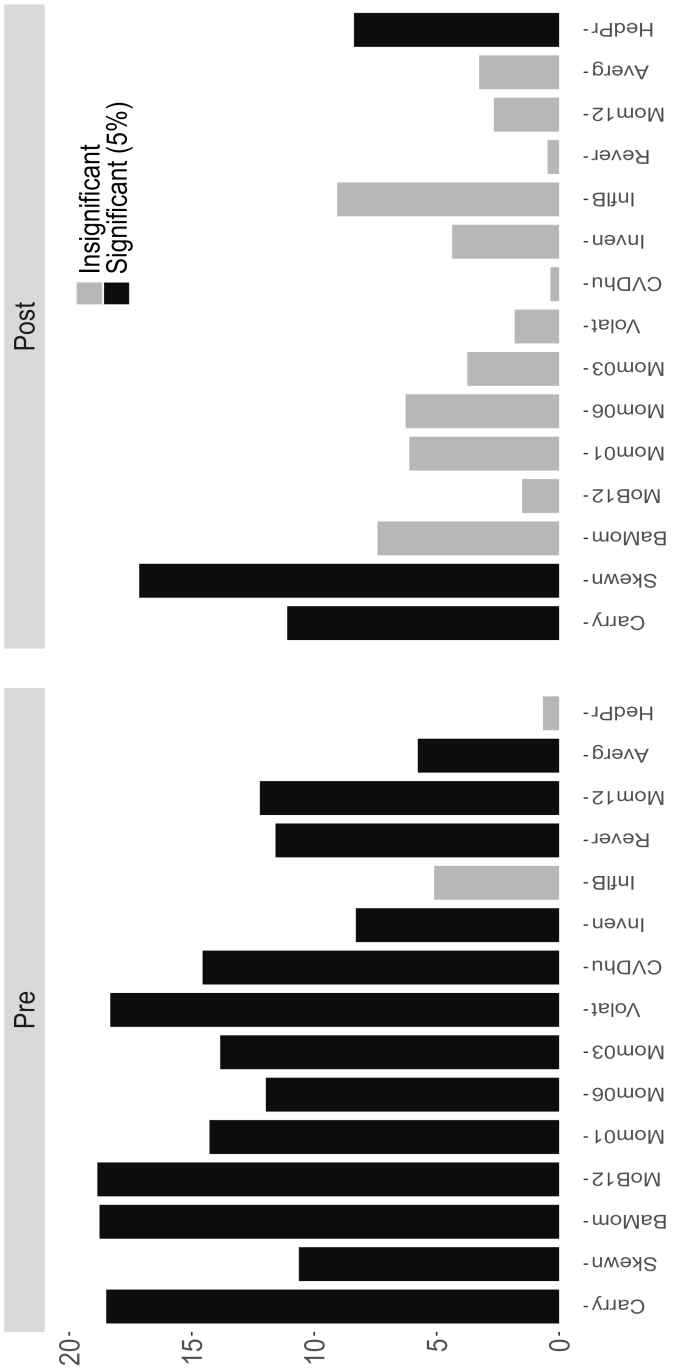

Figure A.2: Average returns to the commodity investment strategies (Pre and Post-Financialization)

This histogram shows the average returns to each of the commodity trading strategies. The returns are annualized excess returns in percentage points. The sample is monthly from March 1986 to August 2021. The left and right panels report the average the returns to the strategies, respectively, over the pre- and post-financialization periods (where the sample is split around January 2004). The choice of 2004 as date for the financialization of commodity markets is driven by the previous literature (see Boons et al. (2014) and Basak and Pavlova (2016), among others). Black and dark grey bars represent, respectively, strategies that deliver returns significant at the 5% and 10% significant level; while, light grey bars represent strategies that deliver average returns not statistically significant. The commodity investment strategies are described in Section A.2 of the Appendix and their performance across the different subsamples (Full Sample, Pre- and Post-Financialization periods) are analyzed in Table 1.

A.4 Styled Fact - Robustness (Date Before Financialization of Commodity Markets)

Table A.2: Descriptive Statistics - Returns to the Commodity Investment Strategies (Over Subsamples, Split around 1999)

This table reports the returns to the commodity investment strategies built on the characteristics described in Section 1.2. The construction of the excess returns takes into account the first notice day convention following Bakshi et al. (2019). We build end-of-month series for commodity returns from March 1986 to August 2021. These data are collected from CRB, Datastream, and Factset. The left panel reports the statistics for the returns to the strategies over the full sample of data (03/1986 to 08/2021). While, the middle and right panels report the descriptive statistics for the strategies, respectively, over the periods before and after 1999, where the sample is split around January 1999. Average returns (*Mean*) and standard deviations (*Std*) are annualized in percentage. *SR* refers to sharpe ratio. We compute test statistics (*tstat*) using Newey and West (1987) corrected standard errors (with lag selection following Andrews (1991)). We denote with ***, **, * estimates significant at the, respectively, 1%, 5% and 10% level.

Full Sample					Pre-1999					Post-1999				
Factor	Mean%	Std%	SR	tstat	Factor	Mean%	Std%	SR	tstat	Factor	Mean%	Std%	SR	tstat
Carry***	14.82	18.76	0.79	4.81	Carry***	14.60	20.24	0.72	2.76	Carry***	14.95	17.90	0.83	3.82
Skewn***	13.86	20.83	0.67	4.00	Skewn*	10.60	21.53	0.49	1.80	Skewn***	15.72	20.45	0.77	3.62
BaMom***	13.13	25.29	0.52	2.84	BaMom***	18.68	23.34	0.80	3.19	BaMom	9.98	26.33	0.38	1.55
Inven**	-7.19	13.93	-0.52	-2.55	Inven	-5.60	14.52	-0.39	-1.34	Inven**	-8.90	13.30	-0.67	-2.30
CVDhu**	7.51	17.34	0.43	2.54	CVDhu***	11.82	16.36	0.72	3.02	CVDhu	5.05	17.87	0.28	1.27
Mom06**	9.14	23.97	0.38	2.48	Mom06	5.42	24.08	0.23	0.89	Mom06**	11.26	23.93	0.47	2.43
Mom01**	10.23	25.05	0.41	2.39	Mom01***	20.55	25.66	0.80	3.16	Mom01	4.36	24.58	0.18	0.83
Volat**	8.35	23.43	0.36	2.34	Volat***	17.32	26.43	0.66	2.97	Volat	3.23	21.44	0.15	0.73
MoB12**	10.24	28.67	0.36	2.16	MoB12*	14.70	29.34	0.50	1.85	MoB12	7.71	28.32	0.27	1.31
Mom03**	8.83	24.41	0.36	2.08	Mom03	11.77	24.30	0.48	1.54	Mom03	7.16	24.51	0.29	1.45
InflB*	7.07	21.80	0.32	1.85	InflB	1.41	19.01	0.07	0.27	InflB**	10.28	23.22	0.44	2.03
Averg*	4.53	12.00	0.38	1.81	Averg*	4.30	9.45	0.45	1.67	Averg	4.66	13.25	0.35	1.31
Rever*	6.07	22.32	0.27	1.68	Rever***	18.38	21.64	0.85	3.50	Rever	-0.93	22.48	-0.04	-0.20
DollB	-4.67	19.71	-0.24	-1.41	DollB	4.61	17.47	0.26	0.94	DollB**	-9.94	20.76	-0.48	-2.31
Mom12	4.83	21.37	0.23	1.38	Mom12**	9.71	20.10	0.48	2.03	Mom12	2.05	22.06	0.09	0.43
HedPr	3.83	18.17	0.21	1.26	HedPr	-2.25	19.53	-0.12	-0.39	HedPr**	7.28	17.30	0.42	2.23
CVSzy	3.78	17.55	0.22	1.18	CVSzy	-0.39	18.45	-0.02	-0.07	CVSzy	6.15	17.01	0.36	1.60
Basis	2.57	25.74	0.10	0.58	Basis	12.57	27.03	0.46	1.63	Basis	-3.12	24.87	-0.13	-0.60
OpeIn	-1.33	18.20	-0.07	-0.43	OpeIn	-3.56	20.34	-0.18	-0.61	OpeIn	-0.07	16.89	0.00	-0.02
Liqui	-0.60	15.57	-0.04	-0.20	Liqui	0.43	15.22	0.03	0.08	Liqui	-1.18	15.79	-0.07	-0.34
Value	0.59	18.14	0.03	0.18	Value	6.48	15.65	0.41	1.55	Value	-2.75	19.37	-0.14	-0.63

A.5 Unconditional Asset Pricing Test - 1st Stage Regressions

Table A.3: Unconditional Asset Pricing Tests (Over Subsamples) - First Stage Regressions

This table shows the regression results from the first-stage of the Fama-MacBeth regressions in Table 2. Hence, we here report OLS estimates of contemporaneous time-series regression of the strategies on the latent risk factors (and a constant). In the asset pricing tests, we use as test assets the returns to the commodity investment strategies presented in Table 1; while, the six candidate factors are the six RP-PCs extracted as in Lettau and Pelger (2020). The left panel reports the results for the test conducted over the full sample of data (03/1986 to 08/2021). While, the middle and right panels report the results for the tests conducted, respectively, over the pre- and post-financialization periods, where the sample is split around January 2004. The choice of 2004 as date for the financialization of commodity markets is driven by the previous literature (see Boons et al. (2014) and Basak and Pavlova (2016), among others). Alphas and R2 are expressed in percentage per month. Standard errors are computed with Newey and West (1987) correction (with Andrews (1991) lag selection). We denote with ***, **, * estimates significant at the, respectively, 1%, 5% and 10% level. The data contain monthly series from March 1986 to August 2021.

	Full-Sample								Pre-Financialization								Post-Financialization							
Variable	$Cons_\alpha$	$PC1_\beta$	$PC2_\beta$	$PC3_\beta$	$PC4_\beta$	$PC5_\beta$	$PC6_\beta$	R2	$Cons_\alpha$	$PC1_\beta$	$PC2_\beta$	$PC3_\beta$	$PC4_\beta$	$PC5_\beta$	$PC6_\beta$	R2	$Cons_\alpha$	$PC1_\beta$	$PC2_\beta$	$PC3_\beta$	$PC4_\beta$	$PC5_\beta$	$PC6_\beta$	R2
Averg	0.01	0.08***	0.19***	-0.05**	-0.12***	-0.02	-0.13***	44.46	0.12	0.05***	0.13***	0.00	-0.09***	-0.03	-0.13***	40.89	-0.09	0.14***	0.27***	-0.03	-0.13***	-0.01	-0.15***	52.06
Carry	0.21	0.26***	0.25***	-0.07	0.14***	0.08*	0.11**	52.35	0.23	0.27***	0.36***	-0.13**	0.16***	0.03	0.11**	64.35	0.32	0.22***	0.09*	-0.11*	0.10*	0.14***	0.10**	44.72
Mom06	0.08	0.36***	-0.17***	0.05**	-0.12***	-0.11***	0.02	70.89	-0.12	0.37***	-0.19***	0.04	-0.13***	-0.11***	0.05	76.43	0.27	0.36***	-0.11***	0.10**	-0.11***	-0.11**	-0.01	64.26
Mom12	-0.07	0.30***	-0.08***	-0.02	0.13***	-0.36***	-0.18***	83.58	-0.04	0.29***	-0.10***	0.05	0.15***	-0.34***	-0.18***	86.24	-0.09	0.28***	-0.10***	-0.10***	0.13***	-0.39***	-0.16***	81.96
InflB	0.07	0.13***	0.30***	-0.34***	-0.16***	0.02	-0.14***	53.31	-0.11	0.13***	0.27***	-0.26***	-0.09	-0.01	-0.04	32.32	0.36	0.14***	0.28***	-0.40***	-0.22***	0.09**	-0.26***	73.83
DollB	-0.02	-0.06**	-0.23***	0.32***	0.08	-0.21***	-0.02	43.99	-0.21	-0.03	-0.11**	0.27***	0.06	-0.24***	0.00	30.08	0.28	-0.11***	-0.40***	0.26***	0.07	-0.16***	-0.06	62.41
CVSzy	0.24	0.10***	-0.18***	-0.05	-0.11***	0.19***	0.01	33.07	0.54	0.11***	-0.25***	-0.05	-0.10*	0.13***	0.01	39.63	0.01	0.10***	-0.09**	-0.01	-0.13***	0.24***	0.03	28.83
CVDhu	0.12	0.14***	0.12***	0.04	-0.10***	-0.14***	0.04	24.07	0.54*	0.15***	0.12***	-0.02	-0.15***	-0.15***	0.01	34.90	-0.37	0.12***	0.16***	0.10*	-0.06*	-0.15***	0.11	16.96
HedPr	-0.05	0.08***	0.08*	-0.06	-0.12**	-0.11***	0.38***	34.22	-0.10	0.05**	0.00	0.04	-0.10*	-0.07	0.41***	31.62	0.00	0.13***	0.13**	-0.11	-0.09	-0.15***	0.34***	44.68
BaMom	-0.06	0.30***	0.19***	-0.05*	0.63***	0.40***	-0.06**	87.07	0.10	0.29***	0.16***	-0.03	0.62***	0.40***	-0.06*	86.71	-0.24	0.32***	0.24***	-0.06*	0.65***	0.39***	-0.05*	88.10
MoB12	0.14	0.44***	-0.21***	0.09***	0.14***	-0.32***	-0.21***	88.84	0.10	0.43***	-0.24***	0.10***	0.16***	-0.32***	-0.27***	92.53	0.10	0.47***	-0.16***	0.12***	0.13***	-0.33***	-0.16***	84.05
Basis	0.19	-0.20***	0.27***	0.50***	-0.23***	0.03	-0.16***	73.65	0.31	-0.23***	0.24***	0.53***	-0.21***	-0.01	-0.18***	77.15	0.09	-0.17***	0.29***	0.50***	-0.26***	0.08**	-0.15**	69.54
Volat	-0.15	0.17***	0.42***	0.15***	-0.26***	-0.20***	-0.38***	67.83	-0.04	0.18***	0.41***	0.17***	-0.28***	-0.16***	-0.36***	73.09	-0.31	0.15***	0.42***	0.14*	-0.24***	-0.25***	-0.36***	60.82
Value	0.08	-0.15***	0.12***	0.22***	0.11***	0.24***	-0.02	53.24	-0.03	-0.14***	0.07**	0.28***	0.13***	0.35***	-0.07	65.05	0.06	-0.19***	0.15**	0.17*	0.12***	0.13***	0.07	50.24
Skewn	-0.13	0.17***	0.53***	0.07**	-0.01	-0.10***	0.42***	70.16	-0.46*	0.20***	0.60***	-0.02	-0.06	-0.10***	0.38***	74.32	0.15	0.14***	0.49***	0.11**	0.02	-0.12***	0.47***	69.74
OpeIn	-0.05	0.03**	-0.01	-0.12***	-0.15***	0.27***	-0.47***	60.85	-0.30	0.05**	0.07	-0.18***	-0.20***	0.29***	-0.52***	63.43	0.21	0.00	-0.09***	-0.11***	-0.12***	0.25***	-0.42***	63.71
Liqui	-0.05	0.03***	0.02	-0.18***	-0.11***	0.24***	-0.29***	52.34	-0.14	0.04***	0.06**	-0.21***	-0.14***	0.23***	-0.29***	54.92	0.06	0.03	0.00	-0.16***	-0.09***	0.26***	-0.31***	50.96
Mom03	-0.14	0.37***	-0.17***	0.06*	-0.30***	0.24***	0.11***	81.49	-0.11	0.37***	-0.11***	0.03	-0.30***	0.29***	0.09**	83.74	-0.13	0.37***	-0.21***	0.07**	-0.31***	0.18***	0.12***	80.59
Mom01	-0.05	0.31***	-0.16***	0.23***	-0.35***	0.39***	0.24***	79.55	-0.01	0.30***	-0.18***	0.28***	-0.35***	0.38***	0.24***	79.17	-0.10	0.30***	-0.19***	0.16***	-0.35***	0.41***	0.24***	80.82
Rever	-0.12	0.11***	0.04	0.57***	0.25***	0.14***	-0.10**	67.83	-0.27	0.11***	0.02	0.59***	0.26***	0.13***	0.12	68.50	0.00	0.14***	0.07*	0.60***	0.24***	0.17***	-0.22***	70.09

A.6 Total Open Interest in Commodity Markets

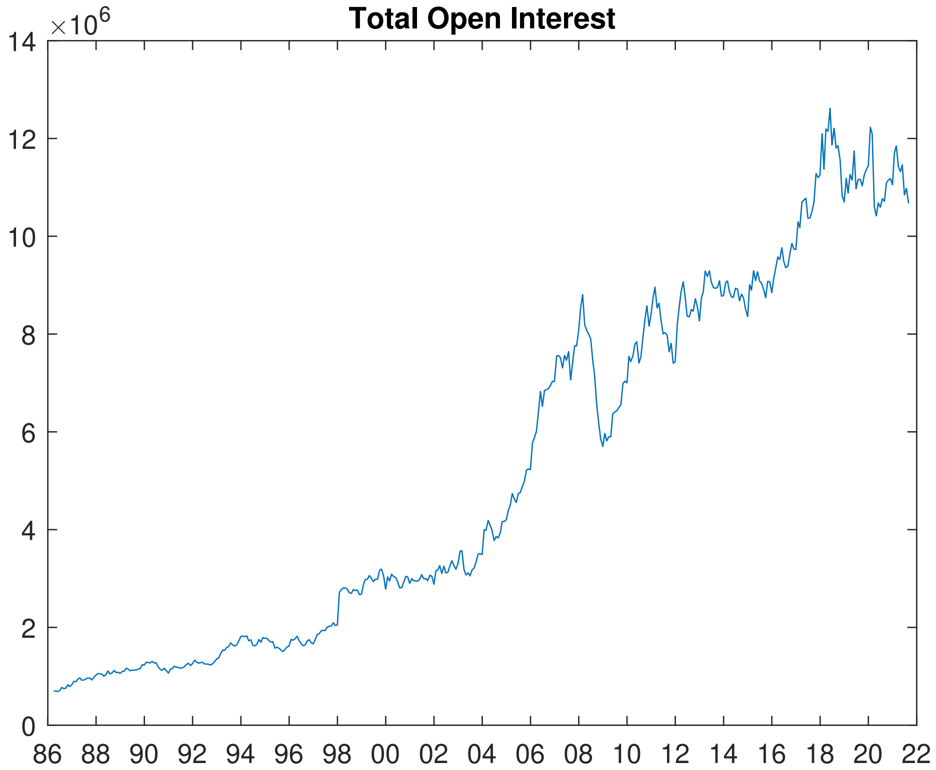

Figure A.3: Total Open Interest
This figure plots the sum of open interest over time. The data cover the period 03/1986-08/2021 and are retrieved from the CFTC.

A.7 Unconditional Asset Pricing Test - Higher Number Of Latent Factors

Table A.4: Unconditional Asset Pricing Tests (Over Subsamples) - 8 Latent-Factor Model

This table reports the results for the second (cross-sectional) stage of the Fama and MacBeth (1973) asset pricing tests with a higher number of factors with respect to Table 2. We use as test assets the returns to the commodity investment strategies presented in Table 1; while, the eighth candidate factors are the first eighth RP-PCs extracted as in Lettau and Pelger (2020). *Panel A* reports the results for the test conducted over the full sample of data (03/1986 to 08/2021); while, *Panel B* and *Panel C* report the results for the tests conducted, respectively, over the pre- and post-financialization periods, where the sample is split around January 2004. The choice of 2004 as date for the financialization of commodity markets is driven by the previous literature (see Boons et al. (2014) and Basak and Pavlova (2016), among others). Mean (*Mean*) and prices of risk (*RP*) for each latent factor (the RP-PCs), as well as for the estimated intercepts, are reported in annualized percentage points. The test-statistics are computed using Newey and West (1987)- ($tstat_{nw}$) and Shanken (1992)-corrected ($tstat_{sh}$) standard errors. Cross-sectional R^2 are in percentage points. The risk premium parameter of the Lettau and Pelger (2020) procedure is set equal to 10. The sample is monthly from March 1986 to August 2021.

Panel A					Full Sample				
Factors	Intercept	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8
Mean (%)		30.77	15.71	5.76	0.39	4.26	4.37	0.68	0.39
RP (%)	0.76	28.18	14.22	4.50	0.88	3.53	4.74	0.28	0.50
$tstat_{nw}$	[0.73]	[3.43]	[3.27]	[0.81]	[0.20]	[0.85]	[1.07]	[0.07]	[0.17]
$tstat_{sh}$	[0.74]	[3.21]	[2.76]	[0.88]	[0.19]	[0.84]	[1.18]	[0.08]	[0.15]
R2 (%)		34.08	79.20	84.95	84.92	88.04	91.25	91.30	91.32

Panel B					Pre-Financialization				
Factors	Intercept	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8
Mean (%)		46.09	15.67	14.51	0.99	2.25	-5.55	-4.38	0.96
RP (%)	0.59	43.63	13.33	13.39	1.71	0.72	-5.97	-3.62	0.42
$tstat_{nw}$	[0.55]	[3.55]	[2.70]	[1.68]	[0.38]	[0.10]	[-1.08]	[-0.65]	[0.09]
$tstat_{sh}$	[0.40]	[3.21]	[1.79]	[1.83]	[0.27]	[0.12]	[-1.02]	[-0.73]	[0.09]
R2 (%)		39.43	58.75	79.43	79.58	79.68	83.82	85.02	85.03

Panel C					Post-Financialization				
Factors	Intercept	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8
Mean (%)		15.88	15.30	-3.29	-0.42	5.96	14.26	6.14	-0.28
RP (%)	1.00	12.72	13.34	-4.88	-0.02	5.53	14.29	6.19	-0.84
$tstat_{nw}$	[0.76]	[1.59]	[1.96]	[-0.76]	[0.00]	[1.16]	[2.55]	[1.07]	[-0.22]
$tstat_{sh}$	[0.76]	[1.18]	[1.89]	[-0.70]	[0.00]	[0.94]	[2.61]	[1.16]	[-0.19]
R2		6.79	41.85	43.87	43.87	49.51	80.75	88.34	88.53

A.8 Macro-Financial Variables and Broader Discussion on the Characterization of the Latent Factors

A.8.1 Broader Discussion on the Characterization of the Latent Factors

This subsection extends the characterization of the latent factors conducted in Section 3.2.

First of all, as we can observe in Table 2, it is important to remind that PCs 4 and 5 neither explain much of the variation in average returns (as, partly, also PC3) nor experience a significant change in their dynamics across the financialization (as also PC2).

However, it is worth highlighting that PC4 is quite neatly identifiable as a basis-momentum factor; and PC5 appears to load primarily on momentum factors, and secondarily on open interest and liquidity.

Regarding the relation of the latent factors with macro-financial risks, as we mention in Section **??**, variations in PC2 partly come from financial variables that can loosely be linked to variations in discount rates. Specifically, this financial variables are mostly interest rates and interest rate differentials, such as the 10 year rate (GS10), the ted spread (ted) and the default yield spread (BAAMAAA)). The intermediary capital variable (icap) also plays a role for PC2, meaning that shocks to leverage in the economy affect returns to commodity mostly through PC2; but also measures that co-move with the global financial cycle (gfc) and global economic conditions (gecon) drive variations in this latent factor. PC5, instead, looks similar to PC3 (which, in turn, we carefully describe in the main part of the paper); whereas PC4 seems to be significantly driven only by shocks to inflation.

Eventually, looking at the results for the pre- and post-financialization periods in Tables A.6 and A.7, a compelling result emerges. Specifically, the finding that variations in PC1 and PC2 over the full sample are associated with variations, respectively, in fx and in discount rates, as well as the finding on the behaviour of PC3, seem to be driven mainly by shocks coming from the post-financialization period. Overall, looking also at the estimates for inflation and volatilities risks, it appears that variations in macro-financial risks tend to be a more relevant determinant of variation in the latent risk factors since the occurrence of the financialization.

A.8.2 Macro-Financial Variables - Tables

Table A.5: Description of the Macro-Financial Variables

This table reports, in the first column (*Variables*) the names of the macro-financial variables used in in Table 4. The second column (*Category*) reports to which broad category the variables belong to. The third column (*Description*) reports brief descriptions of the variables and the sources from which the data are retrieved. Eventually, the fourth (*Start*) and fifth columns (*End*) report the period for which the macro-financial variables are available. For each variable, innovations to the factor are estimated as the first difference of the factor, as the residuals from an AR(1) fitted to the factor itself, or following the FRED-MD dataset of McCracken and Ng (2016) (https://research.stlouisfed.org/econ/mccracken/fred-databases/).

Variables	Category	Description	Start	End
sliq	Financial	Systematic, low frequency, FX (il)liquidity factor. Source: Angelo Ranaldo's website.	01/1991	07/2019
icap	Financial	Intermediary capital risk factor. Source: Zhiguo He's website.	01/1986	06/2021
gfc	Financial	Global financial cycle factor. Source: http://silviamirandaagrippino.com/code-data	01/1986	04/2019
ted	Financial	TED spread. Source: https://fred.stlouisfed.org/series/TEDRATE	01/1986	06/2021
GS10	Financial	10-Year Treasury Rate. Source: https://research.stlouisfed.org/econ/mccracken/fred-databases/	01/1986	04/2021
TWEXAFE	Financial	Trade Weighted U.S. Dollar Index. Source: https://research.stlouisfed.org/econ/mccracken/fred-databases/	01/1986	04/2021
BAAMAAA	Financial	Baa Minus Aaa Corporate Bond Yield. Source: https://research.stlouisfed.org/econ/mccracken/fred-databases/	01/1986	04/2021
S&P 500	Financial	S&P 500. Source: https://research.stlouisfed.org/econ/mccracken/fred-databases/	01/1986	04/2021
equ_vol	Financial	Equity Volatility. Source: constructed as in Bakshi et al. (2019).	01/1986	04/2021
comm_vol	Financial	Commodity Volatility. Source: constructed as in Bakshi et al. (2019).	01/1986	04/2021
gecon	Macro	Global Economic Conditions Indicator. Source: https://sites.google.com/site/cjsbaumeister/research	02/1986	09/2021
INDPRO	Macro	IP Index. Source: https://research.stlouisfed.org/econ/mccracken/fred-databases/	01/1986	04/2021
WPSFD49207	Macro	PPI: Finished Goods. Source: https://research.stlouisfed.org/econ/mccracken/fred-databases/	01/1986	04/2021
CPIAUCSL	Macro	CPI : All Items. Source: https://research.stlouisfed.org/econ/mccracken/fred-databases/	01/1986	04/2021

Table A.6: Latent Factors and Their Relations to Macro-Financial Variables (Pre-Financialization)

This table shows the regression results based on univariate regressions of each latent factor (i.e. the RP-PCs extracted as in Lettau and Pelger (2020)) on a constant and a macro-financial variable (i.e. $RPPC_t^j = \phi_0 + \phi_1 \Delta MacroFin_t^i + \epsilon_t$). In the table, *Cons* refers to ϕ_0, while *Slope* to ϕ_1. The column *R2* reports the R^2 in percentage. We build end-of-month series for commodity returns from March 1986 to August 2021. The sample is split around January 2004. The choice of 2004 as date for the financialization of commodity markets is driven by the previous literature (see Boons et al. (2014) and Basak and Pavlova (2016), among others). Standard errors are computed with Newey and West (1987) adjustment (number of lags selected as in Andrews (1991)). We denote with ***, **, * estimates significant at the, respectively, 1%, 5% and 10% level. Table A.5 in the separate Appendix reports the description of each variable. Results for the subset of variables that are found to be not significant are not reported, but are available upon request.

		PC1			PC2			PC3			PC4			PC5			PC6		
Variables	Category	Cons	Slope	R2	Cons	Slope	R2	Cons	Slope	R2	Cons	Slope	R2	Cons	Slope	R2	Cons	Slope	R2
sliq	Financial	0.040***	-0.027	0.48%	0.013***	0.027**	2.45%	0.008	-0.007	0.13%	0.000	0.015	0.62%	-0.001	-0.004	0.06%	-0.003	-0.025**	2.74%
icap	Financial	0.040***	-0.242	1.09%	0.013**	0.043	0.11%	0.012**	0.072	0.33%	0.000	0.158**	2.04%	0.003	-0.125*	1.36%	-0.005	0.055	0.29%
gfc	Financial	0.039***	-0.087	1.09%	0.013**	0.060**	1.66%	0.012**	-0.028	0.37%	0.001	0.037	0.89%	0.002	-0.038	0.99%	-0.005	-0.015	0.16%
ted	Financial	0.038***	0.012	0.03%	0.013**	0.003	0.01%	0.012**	0.008	0.04%	0.001	-0.033	0.91%	0.002	0.054**	2.61%	-0.005	-0.056**	3.06%
GS10	Financial	0.038***	-0.039	0.39%	0.014**	0.041**	1.37%	0.012**	0.005	0.02%	0.000	-0.026	0.77%	0.003	0.040**	1.90%	-0.005	-0.010	0.12%
S&P 500	Financial	0.042***	-0.467	1.13%	0.011**	0.236*	0.94%	0.011*	0.150	0.39%	-0.001	0.259*	1.54%	0.004	-0.350**	2.98%	-0.005	0.022	0.01%
equ_vol	Financial	0.038***	-7.388**	2.10%	0.013**	1.698	0.36%	0.012**	3.653**	1.72%	0.001	-1.737	0.51%	0.002	2.913**	1.54%	-0.005	-1.012	0.20%
comm_vol	Financial	0.038***	5.405	0.34%	0.013**	-0.530	0.01%	0.012**	8.131	2.55%	0.001	-6.778**	2.33%	0.002	2.521	0.34%	-0.005	-0.595	0.02%
gecon	Macro	0.040***	-0.021	0.19%	0.010*	0.035**	1.64%	0.012*	0.005	0.03%	0.001	-0.001	0.00%	0.005	-0.040**	3.14%	-0.006	0.017	0.58%
WPSFD49207	Macro	0.038***	2.918	0.98%	0.013**	1.167	0.51%	0.012**	-2.363**	2.16%	0.001	1.157	0.68%	0.002	0.112	0.01%	-0.005	-1.740**	1.79%
CPIAUCSL	Macro	0.038***	2.663	0.13%	0.013**	-0.449	0.01%	0.012**	-3.796*	0.90%	0.001	-1.342	0.15%	0.002	-0.167	0.00%	-0.005	0.180	0.00%

Table A.7: Latent Factors and Their Relations to Macro-Financial Variables (Post-Financialization)

This table shows the regression results based on univariate regressions of each latent factor (i.e. the RP-PCs extracted as in Lettau and Pelger (2020)) on a constant and a macro-financial variable (i.e. $RPPC_t^j = \phi_0 + \phi_1 \Delta MacroFin_t^i + \epsilon_t$). In the table, *Cons* refers to ϕ_0, while *Slope* to ϕ_1. The column *R2* reports the R^2 in percentage. We build end-of-month series for commodity returns from March 1986 to August 2021. The sample is split around January 2004. The choice of 2004 as date for the financialization of commodity markets is driven by the previous literature (see Boons et al. (2014) and Basak and Pavlova (2016), among others). Standard errors are computed with Newey and West (1987) adjustment (number of lags selected as in Andrews (1991)). We denote with ***, **, * estimates significant at the, respectively, 1%, 5% and 10% level. Table A.5 in the separate Appendix reports the description of each variable. Results for the subset of variables that are found to be not significant are not reported, but are available upon request.

		PC1			PC2			PC3			PC4			PC5			PC6		
Variables	Category	Cons	Slope	R2	Cons	Slope	R2	Cons	Slope	R2	Cons	Slope	R2	Cons	Slope	R2	Cons	Slope	R2
sliq	Financial	0.015**	-0.062**	3.20%	0.012**	-0.060***	6.88%	-0.005	0.035**	2.05%	0.006	0.018	0.60%	0.006	0.021	1.08%	0.009*	0.015	0.56%
icap	Financial	0.014*	-0.054	0.09%	0.014***	0.405***	10.96%	-0.003	-0.351***	8.49%	0.001	-0.027	0.05%	0.004	0.006	0.00%	0.012**	-0.127	1.77%
gfc	Financial	0.018**	0.117***	8.03%	0.012***	0.098***	12.66%	-0.006	-0.084***	8.75%	0.005	-0.031	1.24%	0.006	-0.018	0.50%	0.009*	-0.022	0.85%
ted	Financial	0.014*	0.018	0.13%	0.014***	-0.050**	2.22%	-0.004	0.060*	3.21%	0.001	0.024	0.55%	0.004	0.020	0.47%	0.012**	-0.010	0.13%
GS10	Financial	0.015*	0.025	0.17%	0.015***	0.053**	1.69%	-0.005	-0.085**	4.43%	0.001	0.010	0.06%	0.004	0.028	0.66%	0.012**	-0.041	1.61%
TWEXAFE	Financial	0.015**	-1.423***	3.46%	0.014***	-1.113***	4.61%	-0.004	1.641***	10.32%	0.001	0.754*	2.39%	0.003	-0.027	0.00%	0.012*	0.232	0.33%
BAAMAAA	Financial	0.015*	-0.064	0.48%	0.014***	-0.183***	8.72%	-0.004	0.159***	6.77%	0.001	0.054	0.84%	0.003	0.008	0.03%	0.013***	0.074	2.36%
S&P 500	Financial	0.010	0.792**	6.05%	0.011**	0.498***	5.20%	0.000	-0.704***	10.73%	0.002	-0.101	0.24%	0.005	-0.240*	1.71%	0.014**	-0.183	1.15%
equ_vol	Financial	0.011	-4.858*	3.58%	0.014***	-2.176**	1.45%	-0.002	3.601*	4.18%	0.002	0.895	0.28%	0.003	2.090**	1.95%	0.012**	0.640	0.21%
comm_vol	Financial	0.011	-3.107	0.32%	0.013***	-8.753**	5.05%	-0.002	9.928**	6.87%	0.002	2.660	0.54%	0.003	5.537***	2.96%	0.012***	3.938	1.70%
gecon	Macro	0.015**	0.024	1.28%	0.015***	0.031***	4.69%	-0.004	-0.030**	4.51%	0.000	0.009	0.48%	0.005	-0.006	0.24%	0.010**	-0.028***	6.28%
INDPRO	Macro	0.015*	-0.310	0.06%	0.013**	1.129	1.62%	-0.003	-0.573	0.43%	0.001	-0.252	0.09%	0.005	-1.259***	2.87%	0.013***	-0.920	1.78%
WPSFD49207	Macro	0.015*	0.662	0.25%	0.014**	2.163*	5.89%	-0.004	-2.037***	5.38%	0.001	-1.472**	3.08%	0.003	-0.944	1.60%	0.012**	-0.189	0.07%
CPIAUCSL	Macro	0.015*	5.699**	2.20%	0.014***	7.335***	7.91%	-0.004	-8.110***	9.96%	0.001	-5.324***	4.71%	0.003	-3.826***	3.06%	0.012	-3.086*	2.30%