Commodity Returns: Lost in Financialization

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

[1]Kelley School of Business
[2]CUNEF University

January 6, 2025

Abstract

The financialization of commodity markets, characterized by a dramatic increase in institutional investors and index capital flows into the commodity futures market around 2004, has significantly impacted the asset class. This paper investigates the effect of this growth in investment capital on the average returns of popular commodity futures trading strategies over time. Our findings reveal that approximately 80% of commodity futures strategies that generated statistically significant average returns before financialization are no longer profitable. Our results suggest that this decline in strategy returns is primarily driven by an adverse change in the average returns of a few systematically priced factors in the cross-section of commodity futures. Furthermore, we find that commodity strategies with relatively higher exposure to the Dow Jones Commodity Index experience a significant reduction in average returns, providing a possible channel for this observed effect. Robustness tests indicate that the publication of commodity strategies in the academic literature accounts for only about 25% of the observed decrease in commodity futures strategy returns.

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

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

*Baba-Yara: Kelley School of Business, Indiana University, 1309 E. 10th St., Bloomington, USA; fababa@iu.edu. Bondatti: CUNEF University, Calle de Almansa 101, Madrid, Spain; m.bondatti@cunef.edu. The authors thank Daniele d'Arienzo, Andrei S. Gonçalves (discussant), Syed Harun (discussant), Nicholas Hirschey, Tobias Lauter (discussant), Giorgio Ottonello, Melissa Prado, Zhongling Qin (discussant), Otto Randl (discussant), Robinson Reyes Pena (discussant), Alberto Rossi, You-Jia Sun (discussant) and Jiri Woschitz (discussant) for helpful discussions and suggestions. Useful comments from participants to the 2024 Commodity and Energy Markets Association, 2024 Cunef Finance Brownbag, 2023 AEFIN Finance Forum, 2023 European Financial Management Association, 2023 Kelley Finance Brownbag, 2023 Financial Markets and Corporate Governance Conference, 2023 Eastern Finance Association, 2023 Southwestern Finance Association, 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 (grant BD/144901/2019). 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

In this paper, we show that an unprecedented increase in investment capital into the commodity futures market has significantly reduced the average returns of most commodity trading strategies. Since 2004, the financialization of commodity futures markets has emerged as a significant phenomenon characterized by a substantial increase in investment capital. Around this period, both retail and institutional investors sought to diversify their portfolios beyond traditional assets, such as stocks and bonds, by investing in commodity futures (Basak and Pavlova (2016)). This trend led to a significant influx of investment capital directed toward indexing, which involves tracking commodity indexes like the S&P Goldman Sachs Commodity Index (SPGCI) and the Dow Jones Commodity Index (DJCI). This growth in index investment capital has been remarkable, increasing from approximately \$15 billion in 2003 to around \$200 billion in 2008 (CFTC (2008), Henderson et al. (2015)).

Policymakers and academic researchers have extensively studied the financialization of commodity futures markets. Early research predominantly focused on the influence of index capital inflows on commodity prices and volatility. However, more recent literature has expanded its scope to investigate the impact of changing price informativeness on firms with significant economic exposure to commodities due to their production processes. For example, Brogaard et al. (2019) find that financialization has diminished the ability of these firms to extract valuable signals from market prices for decision-making purposes.

In this paper, we investigate the impact of financialization on the average returns of twenty distinct commodity futures strategies. Our findings reveal a considerable decline in the informativeness of commodity prices following the financialization of commodity markets. Before 2004, thirteen of the twenty strategies we examined yielded at least marginally significant returns. However, after financialization, only three strategies continue to produce significant returns.

We put forth and examine two possible explanations for this stylized fact. The first hypothesis contends that many of these strategies were previously able to capitalize on idiosyncratic mispricings across various commodity futures. However, the financialization-driven surge in index capital has led to the elimination of these mispricings. The second hypothesis posits that a few systematic commodity factors, such as risk or behavioral factors, underlie the returns of these strategies (see Kozak et al. (2018)). In this scenario, an increase in index investors could decrease the average returns of these factors, thereby lowering the average returns to commodity futures strategies.[1]

To this end, we propose a six-factor asset pricing model and show that this model explains most of the priced variations across the twenty commodity futures strategies. Our findings suggest that the average return of the individual factors has significantly decreased following the influx of investor capital. Consequently, we attribute the decline in average returns to commodity futures strategies as stemming from the reduction in returns to a handful of factors that drive returns in this market.

The existing literature on commodity futures strategies posits that demand and supply imbalances fundamentally drive the positive average returns of these strategies. These imbalances serve as signals for speculative capital, indicating whether additional capital is needed on the demand or supply side to facilitate market clearing. However, the growing presence of commodity indexers, who are mainly net long and unresponsive to demand imbalances, raises questions about the potential impact on the profitability of commodity strategies (see Keynes (1930), Singleton (2014), Henderson et al. (2015)). Our findings speak to this question. Specifically, we find that strategies with higher exposure to index capital experience significant reductions in profitability.

To address the concern that the decrease in strategy returns may be due to the widespread adoption of these strategies rather than the financialization of the market, we conduct tests similar to those in McLean and Pontiff (2016). Our findings suggest that approximately 75% of the average decline in returns can be attributed to the financialization of commodity markets, while the remaining 25% results from the increased popularity of these strategies.

The results in this paper are important for several reasons. First, we uncover a novel channel through which the financialization of commodity futures has impacted the asset

[1] See Chabakauri and Rytchkov (2021) for a general equilibrium model of this effect.

class. Specifically, we show that the influx of index capital has depressed returns to most commodity strategies. Second, we introduce 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. Lastly, we show that the decline in strategy returns follows from significant exposure to commodities in popular long-only commodity indexes such as the DJCI.

This study contributes to three main strands of the literature. Firstly, it adds to the body of research on cross-sectional and time-series predictability in commodity futures markets. A number of recent studies have identified several variables that successfully predict fluctuations 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)).[2] Our research expands upon this literature by comparing the average returns of these strategies before and after financialization. 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. Using a six-factor linear asset pricing model, we find that the returns to all twenty strategies we study are attributable to systematic variation in this cross-section.

Our second contribution is to the strand of the literature that 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

[2] Section 1.2 and the Appendix (A.2) provide an in-depth explanation of the economic rationale behind each variable and their construction.

indexes increased substantially around 2004. Cheng and Xiong (2014) discuss how this phenomenon can affect 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 the 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 associated with 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 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 several prominent commodity futures strategies have materially declined after the financialization of commodity markets. Our results suggest that the fall in strategy returns has happened 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 finds that returns are on average, 58% lower. Similar results also hold in the forex market (Bartram et al. (2023)). 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 most anomalies replicate. We focus on commodity markets as opposed to equity and forex markets. Additionally, we do the exercise in the spirit of McLean and Pontiff (2016), and Bartram et al. (2023), 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 remainder of the paper is organized as follows. Section 1 presents the data and commodity trading strategies. Section 2 examines the impact of commodity markets financialization on average returns to commodity strategies. Section 3 explains the economic rationale behind the decay and explores an index-investing channel that can help rationalize our findings. Section 5 provides robustness checks. Finally, Section 6 concludes.

1 Data and commodity futures strategies construction

Our sample begins in March 1986 and ends in August 2021. We collect end-of-day data on liquid commodity futures contracts from the Commodity Research Bureau (CRB) for the period spanning March 1986 to December 2014. Extending our analysis to August 2021, we incorporate data from Datastream (now Refinitiv) and Factset. March 1986 is selected as the starting date to ensure access to futures returns data for a sufficient number of commodity strategies across different sorting variables, while also maintaining balance between the pre- and post-financialization periods.[3] Overall, our analysis covers 32 commodity futures contracts across four major sectors: agriculture, livestock, energy, and metals.

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. Specifically, we adopt the approach outlined in Bakshi

[3]At each point in time, there are at least 25 commodities to be allocated into (at most 5) portfolios for each sorting variable. Moreover, the starting date aligns with Szymanowska et al. (2014).

et al. (2019) for constructing 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, guaranteeing that its first notice day occurs *after* the end of month $t+1$. By rolling into the shortest maturity contract before the first notice day for each commodity, we ensure that we are never forced to take physical delivery of a commodity.[4] This convention is broadly consistent with Hong and Yogo (2012), Gorton et al. (2013), among others.

The returns of the long and short commodity futures positions are computed as:

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

where r_t^f represents the interest earned on a fully collateralized futures position, and $F_t^{(1)}$ denotes the price of the next maturity futures contract at the end of month t.[5] 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 presents the descriptive statistics for the individual commodity futures excess returns. For approximately 25 out of the 32 commodities under study, the average return is positive. This indicates that a long-only rolling strategy, as outlined earlier, typically yields a positive return over our sample period. Consistent with findings in the literature (see Bakshi et al. (2019)), individual commodity strategies tend to exhibit high volatility, resulting in Sharpe ratios of these long-only rolling strategies

[4]For further details on the futures return construction, we direct interested readers to Bakshi et al. (2019), and to their Table I in the internet Appendix for details on the first notice day convention.

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

typically below 0.25. Additionally, commodity futures often display positive skewness, as observed in Table A.1, providing suggestive evidence for why it is the preferred asset class of Trend-Following traders. Lastly, all commodities are available for at least two-thirds of the sample period.

1.2 Investment strategies definitions

Several variables documented in the literature have been shown to forecast variations in the cross-section of commodity futures returns. In this section, we elaborate on how we construct these characteristics and subsequently form portfolios by sorting the 32 commodities we analyze based on these characteristics. We adhere to the portfolio sorting scheme proposed in the articles that introduced each predictor.

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

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

Third, we construct the basis-momentum factor (*BaMom*) introduced in Boons and Prado (2019), which integrates both momentum and basis fundamental signals. This factor is associated with the slope and curvature of the commodity futures curve, and it reflects imbalances in the supply and demand of future contracts that arise when the market-clearing ability of speculators and intermediaries is impaired.

Fourth, we incorporate a set of volatility-based measures. The coefficient of variations computed using spot prices (*CVDhu*) is sourced from Dhume (2010), while the one using returns (*CVSzy*) is obtained from Szymanowska et al. (2014)). Additionally, we include a volatility factor (*Volat*) constructed as outlined in Gorton et al. (2013), which the authors demonstrate to be linked to inventory levels.

Fifth, we construct the inventory predictor (*Inven*), inspired by Gorton et al. (2013).[6] The theory of storage (Kaldor (1939), Working (1949), Brennan (1958), and Deaton and Laroque (1996)) relates to the timing option inherent in holding a storable commodity. The act of deferring the consumption of a commodity to the future (i.e., when its supply might be scarce) drives its prices above the value of consuming it immediately and generates the convenience yield of holding the commodity. Therefore, according to this theory, the carry of a storable commodity directly translates into the cost of storing it. Holders of inventories earn a convenience yield that is a decreasing and convex function of inventory levels. Thus, the commodity futures risk premium diminishes with increasing inventories. For a comprehensive review of this topic, we also refer to Cheng and Xiong (2014).

The sixth set of variables is related to the hedging pressure theory of Keynes (1930) and Hicks (1939). Hedgers in the futures market tend to be in a net short position and want to eliminate their risk. To incentivize other market participants to assume this risk, hedgers must offer their counterparties a premium to induce them to take the long position. Hedging pressure (*HedPr*) has been demonstrated 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 shown to similarly predict commodity prices beyond imbalances among hedgers.

The seventh set of variables aims to exploit the cross-sectional heterogeneity in the conditional correlation between commodity 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, aiming

[6] We express our gratitude to Martijn Boons for generously sharing the inventory data with us. The inventory factor covers the period only up to 2011.

to exploit a similar correlation between commodity futures returns and exchange rate risk (commodity future prices are typically denominated in a currency, often the U.S. dollar).

Eighth, we sort portfolios based on a liquidity variable (*Liqui*), specifically the Amivest measure proposed by Amihud et al. (1997), which is inspired by Marshall et al. (2012) and Marshall et al. (2013). Expected commodity excess returns can reflect the contract's liquidity, as liquidity may vary across futures or maturities.

Ninth, we construct a standard value factor (*Value*), following the methodology outlined in Asness et al. (2013) and Baba Yara et al. (2021). The value factor is computed using long-term past returns, building on the well-established literature that identifies correlations between past returns and book-to-market ratios (see De Bondt and Thaler (1985), Daniel et al. (1998), and Gerakos and Linnainmaa (2018)).

Tenth, we construct a skewness factor (*Skewn*) following the approach outlined by Fernandez-Perez et al. (2018). As demonstrated by the authors, in line with predictions from theories on investors' skewness preferences or selective hedging, this factor yields significant returns and is capable of explaining the cross-section of commodity futures returns.

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

2 Financialization and the cross-section of commodity futures returns

In this section, we offer a brief overview of the financialization of commodity futures markets and underscore its impact on various commodity futures strategies.

2.1 Background on financialization

Traditionally, the commodity futures market has been primarily dominated by two main participants: commercial hedgers and noncommercial traders. Commercial hedgers typically consist of primary producers of commodities, such as farmers, and primary users of commodities, such as oil refineries. These participants engage in hedging activities to mitigate the risk associated with spot-price fluctuations affecting their business operations. On the other hand, noncommercial traders include managed money traders such as hedge funds and Commodity Trading Advisors (CTAs). These entities take positions on either the long or short side of the market, aiming to help balance out the demand and supply dynamics between primary commodity producers and users.

Around the turn of the millennium, there was a remarkable surge in fund flows into commodity futures as an asset class. Index investment inflows into this space surged from approximately $20 billion in 2003 to more than $200 billion in 2008 and to about $300 billion in 2010 (CFTC (2008), Irwin and Sanders (2011)). Concurrently, the total U.S. exchange-traded futures and futures options trading volume increased from around 630 million contracts per year in 1998 to about 3.2 billion contracts per year in 2007, with growth observed across all commodities. As documented by Boons et al. (2014) and Brogaard et al. (2019), among others, open interest across numerous commodities remained relatively stable between 2000 and 2003. However, it experienced a significant upsurge after 2004.[7] As highlighted in reports by Domanski and Heath (2007), CFTC (2008), and Irwin and Sanders (2011), the number of contracts outstanding in exchange-traded commodity derivatives (and their dollar values) surged by more than threefold. Concurrently, there was a notable shift in the composition of market participants, with the entry of institutional and index investors (Domanski and Heath (2007), Boons et al. (2014), Irwin and Sanders (2011), Brogaard et al. (2019)). By 2008, approximately 24% of the total net notional value of funds invested in commodity indexes was held by "Index Funds," while approximately 42% was held by "Institutional Investors" (CFTC (2008)). Overall, this shift in market structure, dated to the beginning of 2004 in the literature (see Boons et al. (2014), Basak and Pavlova (2016), Brogaard et al. (2019), among others),

[7] Figure A.6 visually demonstrates this surge in commodity investments, which the literature dates back to the beginning of 2004. The figure illustrates the total open interest in the cross-section of commodities expressed in both the number of outstanding contracts and dollar terms. Post-financialization (i.e., after 2004), total open interest spiked to record-high levels. This supports the literature's consensus in dating the financialization to the early 2004 period (refer to Basak and Pavlova (2016), Brogaard et al. (2019), Goldstein and Yang (2021), among others).

is referred to as the financialization of commodity futures markets.

Academics, regulators, and practitioners continue to investigate the precise effects of financialization on commodity markets. Theoretically, Basak and Pavlova (2016) and Goldstein and Yang (2021) show that the rapid increase in indexing activity should impact commodity prices and volatility. However, the empirical evidence remains mixed. Studies by Stoll and Whaley (2010) and Hamilton and Wu (2015) find no evidence of increased index flows affecting commodity prices or volatility. Conversely, Singleton (2014), Henderson et al. (2015), and Brogaard et al. (2019) present evidence to the contrary.

Furthermore, financialization has influenced the informativeness of commodity prices. Goldstein and Yang (2021) demonstrate that as financialization grows, the noise introduced into commodity markets by financial hedgers becomes predominant, outweighing the positive effect of financial speculators on price efficiency, thus diminishing overall price informativeness. Similarly, Brogaard et al. (2019) observe a reduction in commodity price informativeness in response to commodity financialization, though they attribute this effect to the index investing mechanism. An increasing number of empirical studies across other asset classes suggest that index investing, in general, leads to poorer price informativeness (and higher price volatility), as seen in studies such as Israeli et al. (2017), Ben-David et al. (2018), and Coles et al. (2022).

Our paper contributes to the ongoing debate by introducing a new perspective on how increased indexing in commodity markets has impacted their dynamics. Whereas the traditional debate has primarily focused on the impact of financialization on prices and volatility, our contribution centers on the returns generated by commodity strategies, which can be viewed as compensation for speculators. In response to commodity financialization, where commodity prices become less informative (Brogaard et al. (2019) and Goldstein and Yang (2021)), the signals upon which many investment strategies rely may lose their overall informativeness, potentially compromising the profitability of these strategies. To the best of our knowledge, we are the first to explore the implications of commodity financialization for commodity futures trading strategies. This endeavor is particularly enlightening, given that most indexers who have entered this market since 2004 have done so under the assumption that commodity futures offer a risk premium to which one should unconditionally gain exposure.

2.2 Disentangling commodity futures portfolio returns

As is standard in the literature, we sort commodities based on the characteristics outlined in Section 1.2 and 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 the methodologies outlined in the referenced studies and is further detailed in Section A.2 in the Internet Appendix.

Table 1 presents summary statistics for all the long-short commodity futures strategies under study. Across the entire sample, approximately 50% of the strategies exhibit statistically significant average returns at the 5% level, with an additional 15% showing significance at the 10% level (see Figure A.2). We also observed that the long-short portfolios have similar volatilities for individual commodity futures. However, while Sharpe ratios of less than 0.25 are commonplace for individual commodities, Sharpe ratios for strategies that have at least a marginally significant return are higher than 0.25.[8]

When decomposing the return of the strategies into their pre- and post-financialization components, we observe a striking result. Among all the strategies with a significant average return over the entire sample, only Carry and Skewness remain significant in the post-financialization period (see Figure A.3). In other words, about 80% of the commodity strategies that exhibit significance over the entire sample period 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 does not stem from an increase in volatility but rather from a reduction in average returns. This is an

[8]Figure A.1 shows that the correlations among the strategies are generally not very high, except for some strategy pairs, such as momentum strategies. Additionally, when we conduct spanning tests of the return to each strategy on a constant and another strategy, we find that approximately 55% of the unique pairs deliver a statistically significant constant (see Table A.2). This number is even higher if we exclude all those strategies that never deliver significant returns (e.g., $DollB$ or $CvSzy$). This evidence highlights that the returns from the trading strategies tend to capture fairly different dynamics in the commodity futures markets.

[9]When moving the split date even just five years before the financialization of commodities markets (i.e., to January 1999), there are nearly as many strategies with average returns different from zero in the periods before and after the split date (see Table A.3).

interesting fact to document, considering that the discussion in the literature regarding how financialization has impacted commodity markets has primarily focused on the price and volatility channels. Our findings suggest that the volatility of commodity futures strategies has remained relatively unchanged. However, the average returns they generate have experienced a significant deterioration. It is also worth pointing out that hedging pressure is the only strategy with an insignificant average excess return in the pre-financialization period but a significant return post-financialization. This is interesting because hedging pressure is a strategy with a strong theoretical foundation for why it should earn a risk premium in this asset class.

Therefore, Table 1 presents a novel empirical finding in the commodity futures literature: the returns to several commodity futures strategies fade away post-financialization. Thus, we demonstrate that the financialization of commodity futures markets, characterized by the influx of investment capital into this asset class, has impacted the market through a previously unexplored channel.[10] The sudden increase in passive investment capital into an asset can significantly depress average returns.

This effect of financialization is further illustrated in Figure 1, where we plot the difference in returns to the strategies post-financialization against the returns pre-financialization. The negative effect is evident for most of the strategies individually, with commodity anomalies exhibiting higher returns pre-financialization tending to show more pronounced declines in returns post-financialization.[11]

3 What explains the decay in average returns?

In this section, we shed light on the economic rationale behind the decline observed in average excess returns of commodity futures strategies post-financialization.

[10]A decline in returns is also observed in other asset classes (namely, equity and forex) after the publication of trading strategies in academic literature. Therefore, one might worry that the decline in the returns to the commodity strategy we observe post-financialization might be capturing a similar phenomenon. In Section 5.1, we conduct robustness tests that rule out this potential alternative explanation.

[11]Figure A.5 illustrates the evolution of trading strategies not previously applied in commodities by previous literature. As observed, most strategies in commodity markets experience a reduction in profitability post-financialization. Section A.8 provides further details.

3.1 Systematic or idiosyncratic deterioration?

We begin by introducing and evaluating two main hypotheses aimed at elucidating the decline in returns of the commodity strategies around financialization.

First, it is plausible that the returns generated by the factors were primarily due to idiosyncratic mispricing. Therefore, as more investment capital flowed into the asset class, this mispricing was gradually eliminated. Second, we consider the possibility that a simple linear factor model adequately prices the cross-section of these commodity futures strategies. If this is indeed the case, the decline in returns must originate from more fundamental drivers of variation in the economy. Recent research, such as that by Kozak et al. (2018), suggests that the existence of a limited set of pricing factors implies that the returns to these strategies derive from a small set of underlying primitives. This still leaves the question of whether or not the primitives are driven by risk or behavioral factors.

To disentangle the two hypotheses, we employ a simple linear asset pricing factor model that does not take a stance on the nature of the pricing factors. Using the risk-premium principal component analysis (RP-PCA) technique introduced by Lettau and Pelger (2020), we extract latent factors from the cross-section of commodity futures portfolios. RP-PCA is a generalization of PCA designed to extract latent factors that simultaneously fit both the time series and the cross-section of expected returns. We examine the scree plot presented in Figure 2, which displays the first 15 eigenvalues to identify relevant factors. We observe three dominant factors and three weaker yet potentially significant factors. We select these six factors for further analysis.

To assess which hypotheses the evidence supports, we conduct Fama and MacBeth (1973) tests of the returns to commodity futures strategies on the first six RP-PCs over the entire sample period and two subsamples centered around financialization in 2004. This asset pricing approach operates on the principle that, in the absence of arbitrage opportunities, the Euler equation dictates that risk-adjusted returns on each zero-cost portfolio should average to zero. With excess returns to commodity portfolios denoted as R_t and a stochastic discount factor (SDF) represented as M_t, the following relationship should hold:

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

More specifically, an 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 allows for a beta representation of the form:

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

where the risk premia to a particular strategy i depend on the price of risk (λ) and the strategy's loadings on the factors (β_i).[12] Thus, in our analysis, this corresponds to the regression coefficients of the excess returns to each commodity investment strategy on the (latent) risk factors, i.e., the RP-PCs. We estimate these coefficients using a standard Fama and MacBeth (1973) two-stage procedure, as commonly done in the literature.

Table 2 presents the results of the analysis. In Panel A, it is evident that the six latent factors can explain a substantial portion of 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.[13] This pattern persists when examining Panels B and C, representing the pre-financialization and post-financialization periods, respectively. These findings contradict the first hypothesis, suggesting that the strategies do not represent idiosyncratic mispricing that has been arbitraged away after financialization.

The analysis further reveals that a significant portion of the principal components (PCs) have experienced changes in risk premia post-financialization. Specifically, over 60% of the PCs have seen reductions in risk premia, with PC3 experiencing reductions exceeding 100% and PC2 experiencing reductions as small as 2%. PCs 5 and 6 have instead both experienced an increase. These results provide suggestive evidence in favor

[12]The relationship between the factor prices and the factor loadings is determined by the covariance matrix of the factors (Σ_h): $\lambda = \Sigma_h b$.

[13]Moreover, the alphas in the first step are all jointly equal to zero in the time-series regressions, as highlighted by the results of the GRS test. Similarly, the intercept is statistically zero when we allow for a free intercept in the cross-sectional tests.

of hypothesis two, indicating that the most likely reason for the lack of significant average excess returns in almost all commodity futures strategies post-financialization is associated to adverse changes in the systematic primitives that drive them.[14]

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 for individual asset pricing tests. To address this concern, we report the first stage Fama and MacBeth (1973) results in Table A.4. These results confirm that the six-latent factor model adequately prices all twenty commodity futures factors in both the total sample and both sub-samples.

Figure 3 illustrates the ten-year rolling average of the latent factors, providing insights into their dynamics over time.[15] Notably, the dominant systematic factor (PC1) exhibits a trend of declining average returns. Before financialization, the average return of PC1 consistently exceeded 3%, but it has since trended toward a negative price of risk. In fact, the last ten-year average of this factor is currently negative. Additionally, the third latent factor (PC3) has experienced a significant deterioration in average returns. Lastly, the figure highlights changes in the price risk of PC6, which exhibited a marginally negative average return pre-financialization but has shown a significantly positive average return in the post-financialization period. These observations further underscore the impact of financialization on the dynamics of commodity futures markets.

Taken together, the results from the asset pricing test support the hypothesis that the influx of investment capital into the commodity futures market around 2004 is indeed associated with a deterioration in the returns to the strategies under study. Furthermore, this reduction occurred through the influence of a handful of systematic factors that the strategies load on.[16]

[14]It's worth noting that the results remain consistent even when considering a seven- or eight-latent factor model, as additional RP-PCs exhibit limited explanatory power across all subsamples and lack significant price of risk (see Table A.5).

[15]Using total open interest as a proxy for flows, Figure A.6 shows that the flows of capital into the commodity futures market kept increasing even after the large sudden spike around the financialization.

[16]This finding underscores the significance of these systematic factors in driving variations in the returns of commodity futures strategies. Given that this is a new factor model in the cross-section of commodities, further characterization of these latent factors can provide additional insights into their properties and their role in shaping commodity market dynamics. Thus, in the Appendix A.12, we take the opportunity to characterize some of their properties better.

3.2 Index flow channel

Next, we delve into index investing as a potential channel behind the observed systematic deterioration in average excess returns of commodity futures strategies post-financialization. We utilize data on constituents of the Dow Jones Commodity Index and their weights, sourced from Standard & Poor's, spanning from January 2000 onwards.

To test this potential mechanism, we employ the following baseline specification:

$$R_{i,t+1} = \alpha_i + \beta_1 D_{i,t} + \Gamma D_{i,t} * \delta_t + e_{i,t+1}. \tag{4}$$

where: i) $R_{i,t+1}$ represents the returns to commodity investment strategy i, ii) α_i is a dummy capturing strategy fixed effects, iii) $D_{i,t}$ is a dummy variable that takes the value one if strategy i at time t has exposure to any commodity in the top-3 weighted commodities in the DJCI index[17], 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. This last term, interpreted as an "exposure by time fixed effects" indicator, allows us to control for the increasing inflows of capital into the commodity markets (and, therefore, into the commodities index) over time. This variable also absorbs variation coming from other potential time-varying confounding omitted effects that impact index-exposed and non-index-exposed strategies differently.

The results of these tests are presented in Table 3. The coefficient of interest, β_1, captures the average effect on the returns to the strategies of trading commodities with high weights in the commodity index. As observed in column (1), exposure to the index has a significant negative effect on the average returns of a strategy. Specifically, a commodity investment strategy experiences about a 70 basis points drop in returns whenever it trades commodities with top-3 weight in the DJCI.

Furthermore, column (2) repeats the analysis with the inclusion of an open interest variable as a control, capturing additional omitted effects of the increasing capital inflows

[17]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 four times larger than the average weight of the third highest weighted commodity.

not perfectly captured by the coefficients Γ. Similarly, column (3) includes a dollar open interest variable, i.e., open interest multiplied by the spot price. Overall, the results remain robust to the inclusion of these additional controls.

In summary, the findings indicate that index flow, and therefore index investing, is a fundamental channel through which financialization has impacted commodity futures strategies.

4 Model

In this section, we develop a dynamic general equilibrium model to examine how the introduction of index investing influences investment strategies within the market for risky assets. Building upon the two-tree, two-agent framework of Chabakauri and Rytchkov (2021), our model incorporates index investors and active traders, allowing us to analyze the differential impacts of indexing on asset returns and investment behaviors.

4.1 Model Framework

Our economy consists of two agents, Agent A and Agent B, two Lucas trees representing two risky assets, and two goods. Agent A represents hedge funds and similar market participants who actively trade individual assets. Agent B represents investors who participate in the market exclusively through index investing. The two Lucas trees correspond to two distinct risky assets with differing expected returns and volatilities, facilitating the exploration of heterogeneous investment strategies.

Time is continuous and extends indefinitely, $t \in [0, \infty)$. Uncertainty is modeled via a probability space $(\Omega, \mathcal{F}, \mathbb{F}, P)$, supporting a two-dimensional Brownian motion $\vec{Z} = (Z_1, Z_2)' \in \mathbb{R}^2$. The filtration $\mathbb{F} = (\mathcal{F}_t)_{t \in [0,\infty)}$ is the usual augmentation generated by the Brownian motions, with $\mathcal{F}_\infty = \mathcal{F}$.

4.1.1 Endowments, prices, and assets

Each Lucas tree produces a differentiated good, with output following a geometric Brownian motion. The output of tree j is defined as:

$$dY_{j,t} = Y_{j,t}\left(\mu_{j,t}dt + \sigma'_{j,t}d\vec{Z}_t\right), \quad j \in \{1,2\} \tag{5}$$

where $\mu_{j,t}$ is the drift of the process and $\sigma_{j,t}$ is the diffusion coefficient. We model each tree as being driven by a single Brownian shock and so $\sigma_{j,t}$ has zeros every where besides the $i'th$ index.

The prices of these goods are denoted by $p_{1,t}$ and $p_{2,t}$. P_t^i is the price of the consumption basket for investor i. All prices are defined with respect to a global numeraire taken to be the CES-basket with a on good 1. Specifically, we normalize $[ap_{1,t}^{1-\theta} + (1-a)p_{2,t}^{1-\theta}]^{1/(1-\theta)}$ to 1.

4.1.2 Returns

Both trees are traded equity assets with returns given by:

$$dR_{j,t} = \frac{dQ_{j,t}}{Q_{j,t}} + \frac{p_{j,t}Y_{j,t}}{Q_{j,t}}dt \tag{6}$$

where $Q_{j,t}$ is the price of the tree j at time t. Economically, $\frac{dQ_{j,t}}{Q_{j,t}}$ represents the capital gain and $\frac{p_{j,t}Y_{j,t}}{Q_{j,t}}$ the dividend yield which together give the total return for owning tree j at time t[18].

We redefine the dividend yield as $F_{j,t} = \frac{p_{j,t}Y_{j,t}}{Q_{j,t}}$, allowing us to rewrite Equation 6 as:

$$\begin{aligned} dR_{j,t} &= \frac{d(p_{j,t}Y_{j,t}/F_{j,t})}{(p_{j,t}Y_{j,t}/F_{j,t})} + F_{j,t}dt \\ dR_{j,t} &= \mu_{R_j,t}dt + \sigma_{R_j,t}d\vec{Z}_t \end{aligned} \tag{7}$$

The supply of each equity asset is normalized to unity and there also exists a bond,

[18] See Appendix A.13 for more details on the derivations

which is locally riskless in units of the numeraire. Its price is B_t and the corresponding instantaneous return is r_t, so that $dB_t = B_t r_t dt$.

Stocks constitute a market portfolio (index), which pays the aggregate dividend $Y_t = p_{1,t}Y_{1,t} + p_{2,t}Y_{2,t}$ and has price $Q_t = Q_{1,t} + Q_{2,t}$. We show in appendix the return process for the index can be expressed as:

$$dR_t = \mu_{R_I,t}dt + \sigma'_{R_I,t}d\vec{Z}_t,$$

with

$$\mu_{R_I,t} = s_t\mu_{R_1,t} + (1 - s_t)\mu_{R_2,t}, \quad \sigma_{R_I,t} = s_t\sigma_{R_1,t} + (1 - s_t)\sigma_{R_2,t}.$$

Since $s_t = \frac{Q_{1,t}}{Q_t}$, the drift and volatility of the index return are determined by the individual drifts and volatilities of the component assets, weighted by their relative importance in the total index value. It is important to explicitly state that the index volatility simplifies to a linear combination because the Brownian shocks are orthogonal.

4.1.3 Preferences

Each investor has recursive preferences over consumption a la Duffie and Epstein (1992). Preferences are given for $i \in \{A, B\}$ by:

$$V_{i,t} = \max_{\{C^i_{1,u}, C^i_{2,u}, w^i_{1,u}, w^i_{2,u}\}_{u=t}^{\infty}} \mathbb{E}_t\left[\int_t^{\infty} f^i(C^i_u, V^i_u)du\right] \tag{8}$$

$$f^i(C, V) = \left(\frac{1-\gamma^i}{1-1/\psi^i}\right)V\left[\left(\frac{C}{[(1-\gamma^i)\,V]^{1/(1-\gamma^i)}}\right)^{1-1/\psi^i} - \rho^i\right]$$

where γ^i is the coefficient of relative risk aversion, ψ^i is the elasticity of intertemporal substitution and ρ^i is the discount rate.

The consumption basket is composed of the two goods, which aggregates under constant elasticity of substitution θ and bias in consumption α^i.

$$C^i_t = \left[{\alpha^i}^{\frac{1}{\theta}}{C^i_{1,t}}^{\frac{\theta-1}{\theta}} + (1-\alpha^i)^{\frac{1}{\theta}}{C^i_{2,t}}^{\frac{\theta-1}{\theta}}\right]^{\frac{\theta}{\theta-1}} \tag{9}$$

We introduce this basket to define a real numeraire and simplify relative price analysis.

4.1.4 Investment Strategies and Budget Constraints

Investors allocate their wealth between risky assets and the risk-free bond, but their investment opportunities differ based on their type.

Agent A (Active Traders) can invest directly in the individual risky assets. Let $w_{j,t}^A$ denote the fraction of Agent A's wealth allocated to asset j, for $j \in \{1, 2\}$. The remaining wealth $(1 - w_{1,t}^A - w_{2,t}^A)$ is invested in the risk-free bond. Agent A chooses $w_{1,t}^A$ and $w_{2,t}^A$ to maximize their utility.

The budget constraint for Agent A is:

$$\frac{dW_t^A}{W_t^A} = \left(r_t + \sum_{j=1}^{2} w_{j,t}^A \left(\mu_{R_j,t} - r_t \right) - P_t^A c_t^A \right) dt + \left(\sum_{j=1}^{2} w_{j,t}^A \sigma_{R_j,t} \right)' d\vec{Z}_t \tag{10}$$

Agent B (Index Investors) is restricted to investing only through the market index. Let $\hat{w}_t^B$ denote the fraction of Agent B's wealth allocated to the index. The remaining wealth $(1 - \hat{w}_t^B)$ is invested in the risk-free bond. Agent B cannot adjust the allocations to individual assets independently of the index.

The portfolio holdings for Agent B in individual assets are derived from the index composition. Specifically, the weight on asset j for Agent B is:

$$w_{j,t}^B = \hat{w}_t^B \cdot \frac{Q_{j,t}}{Q_t} \tag{11}$$

where $\hat{w}_t^B$ can be thought of as just leveraging up or down the market index. Thus, the budget constraint for Agent B is:

$$\frac{dW_t^B}{W_t^B} = \left(r_t + \hat{w}_t^B \left(\mu_{R_I,t} - r_t \right) - P_t^B c_t^B \right) dt + \left(w_t^B \sigma_{R_I,t} \right)' d\vec{Z}_t \tag{12}$$

Both agent types use the proceeds to purchase their desired consumption basket $c_t^i = C_t^i / W_t^i$ at price P_t^i. Investors are also subject to the transversality condition, and we also have $W_t^i \geq 0$ and $W_0^i > 0$.

4.2 Equilibrium

An equilibrium is defined as a set of stochastic processes for the risk-free rate r_t, expected excess returns μ_{Qt}, diffusions of returns Σ_{Qt}, consumption streams $C_{jt}, j = A, B$, and portfolio strategies $\omega_{jt}, j = A, B$, such that:

1. **Optimization**: C_{jt} and ω_{jt} solve the utility maximization problem for each agent j.

2. **Market Clearing**: Aggregate consumption equals aggregate dividends: $C_{At}+C_{Bt} = Y_t$.

3. **Asset Market Clearing**: The supply and demand for each asset clear the market:

$$w^A_{1,t}W^A_t + w^B_{1,t}W^B_t = Q_{1,t} \tag{13}$$

$$w^A_{2,t}W^A_t + w^B_{2,t}W^B_t = Q_{2,t} \tag{14}$$

$$(1 - w^A_{1,t} - w^A_{2,t})W^A_t + (1 - w^B_{1,t} - w^B_{2,t})W^B_t = 0 \tag{15}$$

Given that Agent B invests via the index, we have $w^B_{1,t} = Q_{1,t}/(W^A_t + W^B_t)$ and $w^B_{2,t} = Q_{2,t}/(W^A_t + W^B_t)$.

4.2.1 State variables

Wealth share

The first variable that summarises the state of the economy is the wealth share of the first investor. The wealth share is a relevant state-variable in our economy because wealth is not constant as in standard log utility models. Additionally, wealth is not a monotonic function of other fundamentals in the economy. Wealth share fundamentally captures time-varying Negishi weights as discussed in Dumas et. 2000, Anderson (2005) or Colcaito and Croce (2011, 2013).

$$x_{i,t} = \frac{W_t^i}{\sum_{k=0}^{M-1} W_t^k}$$
$$x_{i,t} = \frac{x_{i,t}^1}{x_t^2}$$

We solve for dynamics of the process in A.14 and the dynamics of $x_{i,t}$ follows:

$$\frac{dx_{j,t}}{x_{j,t}} = \left(\mu_{w_i,t} - \mu_{x_t^2,t} + \sigma'_{x_t^2,t}(\sigma_{x_t^2,t} - \sigma_{w_i,t})\right) dt + \left(\sigma_{w_i,t} - \sigma_{x_t^2,t}\right)' d\vec{Z}_t \tag{16}$$
$$\frac{dx_{j,t}}{x_{j,t}} = \mu_{x,t} x_t dt + x_t \sigma'_{x,t} d\vec{Z}_t$$

Realtive supply of trees

The second state variable is the relative supply of good 1. This is defined as:

$$y_t = \frac{Y_{1,t}}{Y_{1,t} + Y_{2,t}} \tag{17}$$

We solve for dynamics of the process in A.14 and the dynamics of $y_{j,t}$ follows:

$$\frac{dy_{j,t}}{y_{j,t}} = \left(\mu_j - \mu_{y_t^2,t} + \sigma'_{y_t^2,t}(\sigma_{y_t^2,t} - \sigma_{Y_j})\right) dt + \left(\sigma_{Y_j} - \sigma_{y_t^2,t}\right)' d\vec{Z}_t \tag{18}$$
$$\frac{dy_{j,t}}{y_{j,t}} = \mu_{y,t} y_t dt + y_t \sigma'_{y,t} d\vec{Z}_t$$

5 Robustness tests

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

5.1 Financialization versus academic research

Thus far, we have argued that financialization is associated with declining returns to commodity futures strategies. However, a plausible alternative narrative is that the publication of these strategies in academic journals may trigger this decline, independent of financialization. This is a compelling alternative because most strategies under examination were published post-financialization. To investigate this possibility, we adopt an approach akin to McLean and Pontiff (2016) in equity and Bartram et al. (2023) in forex. Formally, we run variations of the following baseline specification to assess this alternative hypothesis:

$$R_{i,t} = \alpha_i + \beta_1 \textit{Post-Financialization Dummy}_{i,t} + \beta_2 \textit{Post-Publication Dummy}_{i,t} + e_{i,t}. \quad (19)$$

Table 4 presents the results, with standard errors for all the tests clustered on time. Specifically, in our first test, we regress the returns to all the strategies we study on a dummy variable that takes the value zero before January 2004 (i.e., pre-financialization) and one afterward (i.e., post-financialization). The estimated coefficient on the dummy variable for this specification is negative and statistically significant, confirming the findings from our portfolio sorts that the returns to the trading strategies decay post-financialization. Specifically, starting from 2004, the average returns to a typical commodity futures strategy have declined by about 51 basis points per month. The pre-financialization average returns of the strategies are approximately 78 basis points per month.

In the second test, we regress the returns to the commodity strategies on a dummy that takes the value zero before the publication date of a strategy and one afterward. While previous literature has documented a decrease in returns to trading strategies post-academic publication in equity and forex markets, the dynamics in commodities have not been analyzed yet. The estimated coefficient from this second specification is similarly negative and significant. Hence, the result suggests that the typical commodity strategy has also experienced a return decay after the strategy was published in an academic journal, quantifiable at around 46 basis points per month.

To disentangle which of these two alternatives drives what fraction of the decay, we regress the returns to the strategies on two dummies. The first dummy takes the value of zero before financialization and one afterward. The second takes the value of zero before publication in the academic literature and one afterward. We report the results as specification 3 in the table. 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 the estimate is statistically insignificant. Therefore, the results of this test strongly suggest that the financialization of commodity markets is associated with the most decay in returns rather than 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) confirm the findings from specification (3) regarding the financialization effect.[19]

Overall, around 75% of the observed decay in returns to commodity strategies can be attributed to the financialization of the market, with the remaining portion stemming from the publication of the strategies in the academic literature.

5.2 Potentially confounding periods after the financialization

To address potential concerns regarding the influence of the global financial crisis (GFC) and the COVID-19 pandemic on commodity futures strategy dynamics, we repeat the baseline analysis but exclude the GFC years (2007 to 2009) and the COVID-19 year (2020) from our sample.[20]

Our analysis indicates that the key findings remain qualitatively consistent. As depicted in Table A.11, we continue to observe a decline in the returns to most commodity strategies post-financialization. This evidence corroborates the idea that the return deterioration is associated with systematic dynamics related to the financialization of commodity futures markets. To further reinforce the robustness of our findings, it is worth mentioning that in a specification similar to column (1) of Table 4, the magnitude of the

[19] Additionally, if we regress the returns on the stock market on the financialization dummy, the estimated coefficient is small but positive. However, including the returns to the market as a control in the specifications in Table 4 does not affect the results. This evidence suggests that the behavior of the overall market does not seem to explain our findings. The results of these additional tests are available upon request.

[20] However, it is important to note that by omitting the early years of the financial crisis, we may also remove some dynamics stemming from the late part of the financialization that occurred during that period.

post-financialization dummy remains significant, indicating a decay of approximately 50 basis points per month. In other words, the average returns to a typical commodity futures strategy have deteriorated by about 50 basis points per month after financialization, even when the GFC and COVID periods are removed from our sample.

6 Conclusions

In this paper, we investigate the impact of financialization and the growth of indexing on the average returns to commodity futures investment strategies. We show that of the thirteen commodity futures strategies with a significant average return before the financialization of commodity markets, only three remain profitable afterward. We find that this decline in strategy returns is primarily associated with a dramatic fall in the average returns of the systematic factors driving expected returns in this asset class.

Our findings offer compelling evidence supporting the notion that the financialization of commodity markets and indexing has impacted certain commodity market participants. Our results align with the model of Sockin and Xiong (2015), in which information frictions hinder commodity futures market participants from effectively utilizing price and volume data to make profitable investment choices.

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, Anthony Garratt, and Yan Xu, 2023, Mispricing and risk premia in currency markets, *Journal of Financial and Quantitative Analysis* 1–93.

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 .

Chabakauri, Georgy, and Oleg Rytchkov, 2021, Asset pricing with index investing, *Journal of Financial Economics* 141, 195–216.

Chen, Andrew Y, and Tom Zimmermann, 2020, Publication bias and the cross-section of stock returns, *The Review of Asset Pricing Studies* 10, 249–289.

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* .

Domanski, Dietrich, and Alexandra Heath, 2007, Financial investors and commodity markets, *BIS quarterly review, March* .

Duffie, Darrell, and Larry G Epstein, 1992, Stochastic differential utility, *Econometrica: Journal of the Econometric Society* 353–394.

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.

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.

Tang, Ke, and Wei Xiong, 2012, Index investment and the financialization of commodities, *Financial Analysts Journal* 68, 54–74.

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

7 Tables and Figures

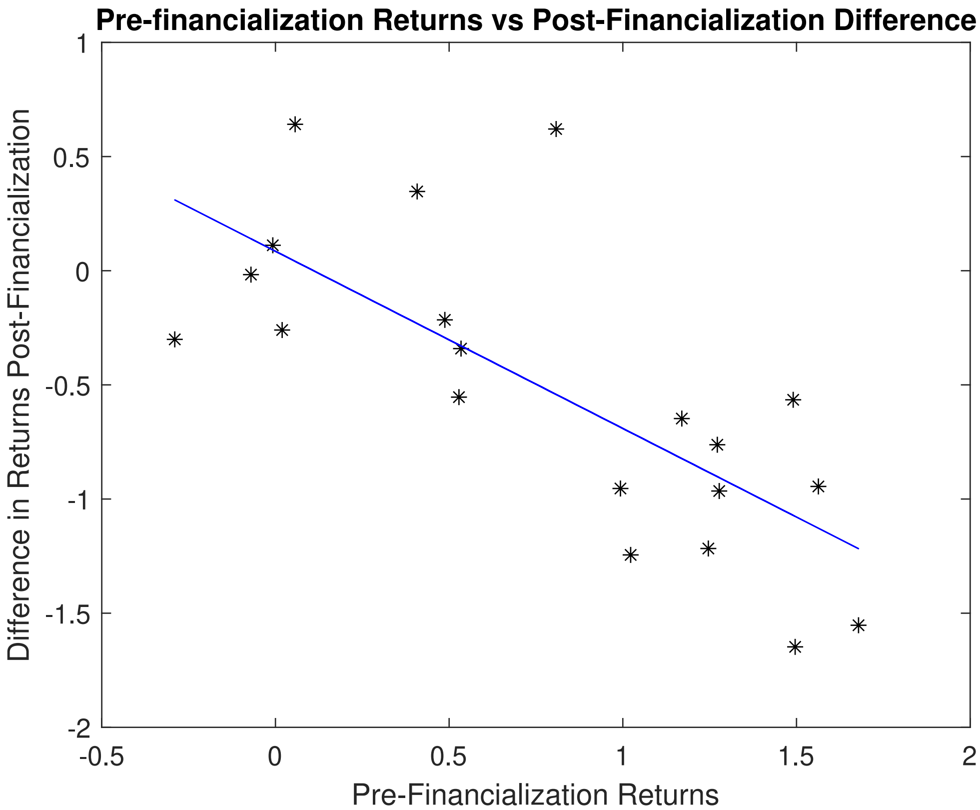


Figure 1: Relation between pre- and post-financializtion returns

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 period covers monthly data 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) is 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 the 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 (i.e., 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. We compute test statistics (*tstat*) using Newey and West (1987) corrected standard errors (with lag selection following Andrews (1991)). Cross-sectional R^2 are in percentage points. Additionally, we report (GRS) F statistic p-values (F p-value) to test the joint null hypothesis that all alphas for the time series regressions are jointly zero, as well as the χ statistic p-values (χ p-value) to test whether, when allowing a free intercept in the OLS cross-sectional (CS) regression, the alphas are zero. allowing a free intercept 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.4 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
F p-value	(0.993)						
χ p-value	(0.996)						

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
F p-value	(0.935)						
χ p-value	(0.984)						

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
F p-value	(0.996)						
χ p-value	(0.999)						

Table 3: 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-3 weighted commodities in the index, and a dummy that at t is equal to one for all strategies that have at least a commodity with top-3 exposure to the index. All regressions include factor fixed effects. Column (1) reports the results when restricting the commodities to have top-3 weights in the DJCI Index over the (monthly) period 01/2004 to 08/2021 (i.e., the post-financialization period). Column (2) repeats the same exercise but adding open interest as a control variable; while, Column (3) adds dollar open interest as a 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)
$D_{i,t}$	-0.707***	-0.568*	-0.549**
Observations	4139	4139	3933
Factor Fixed Effects	Yes	Yes	Yes
Exposure by Time Fixed Effects	Yes	Yes	Yes
Control for Open Interest	No	Yes	No
Control for Dollar Open Interest	No	No	Yes

Table 4: 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

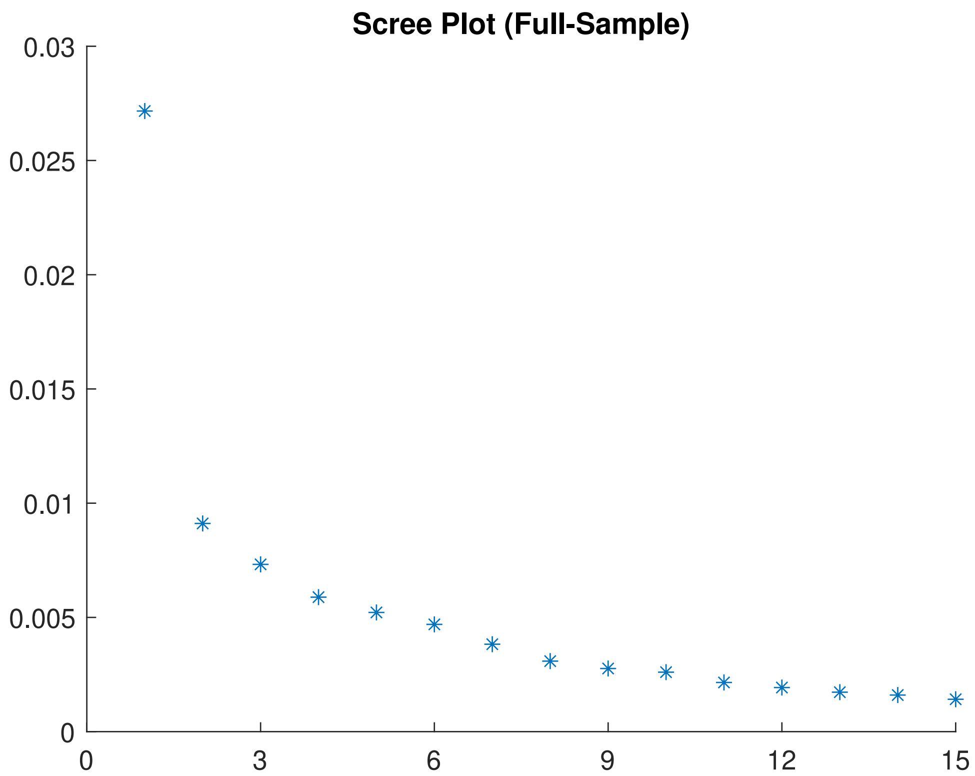


Figure 2: First 15 eigenvalues

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 sample period covers monthly data from March 1986 to August 2021.

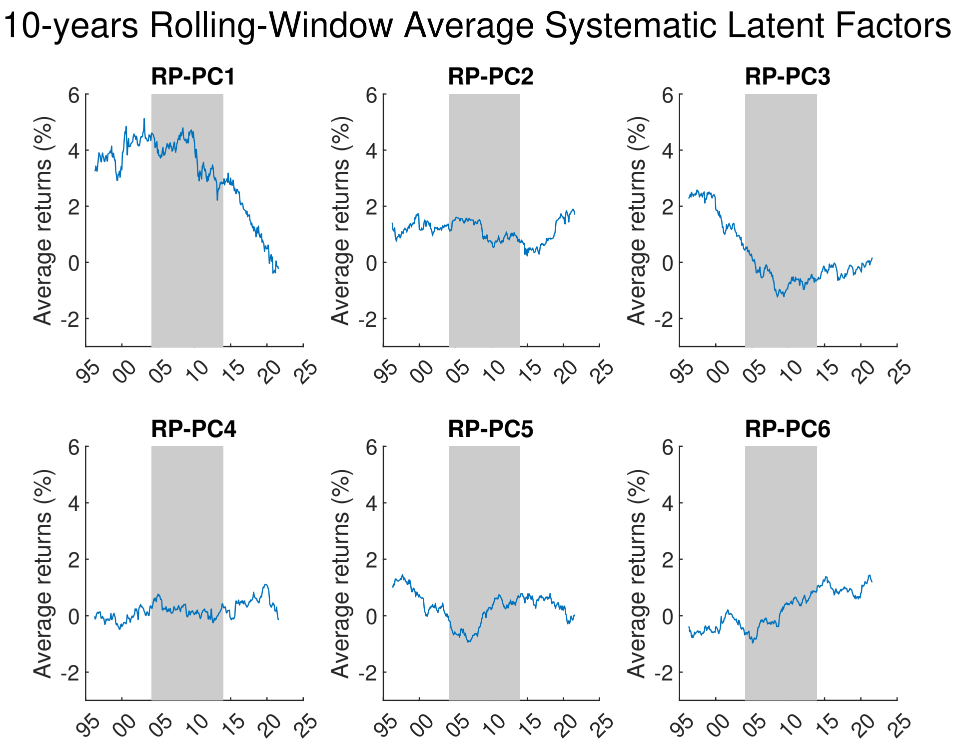


Figure 3: Rolling-window average latent factors

This figure plots the ten-year rolling average return of each RP-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 sample period covers monthly data 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.

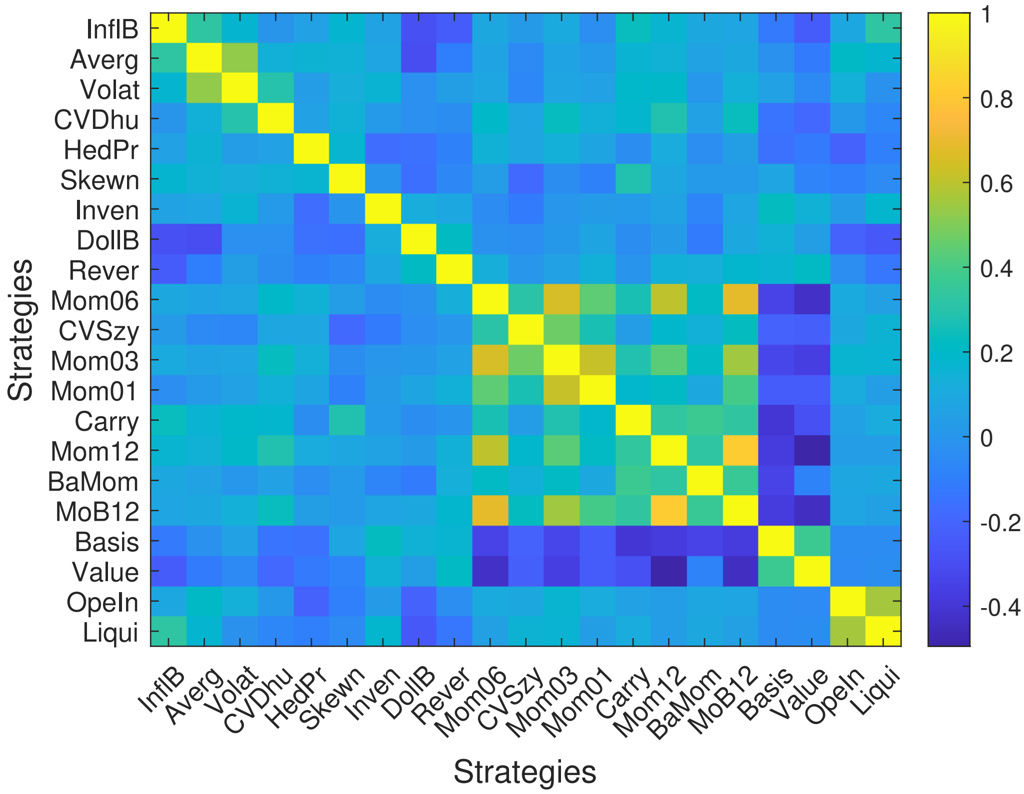


Figure A.1: Correlation matrix of the commodity investment strategies (Full Sample)
This heatmap shows the correlation matrix of the commodity trading strategies. The sample period covers monthly data from March 1986 to August 2021. The commodity investment strategies are described in Section A.2 of the Appendix and their performance across the different subsamples (Full Sample, Pre- and Post-Financialization periods) is analyzed in Table 1.

A.3 Correlation Matrix of the Commodity Trading Strategies

A.4 Average Returns to the Commodity Trading Strategies (Across Subsamples)

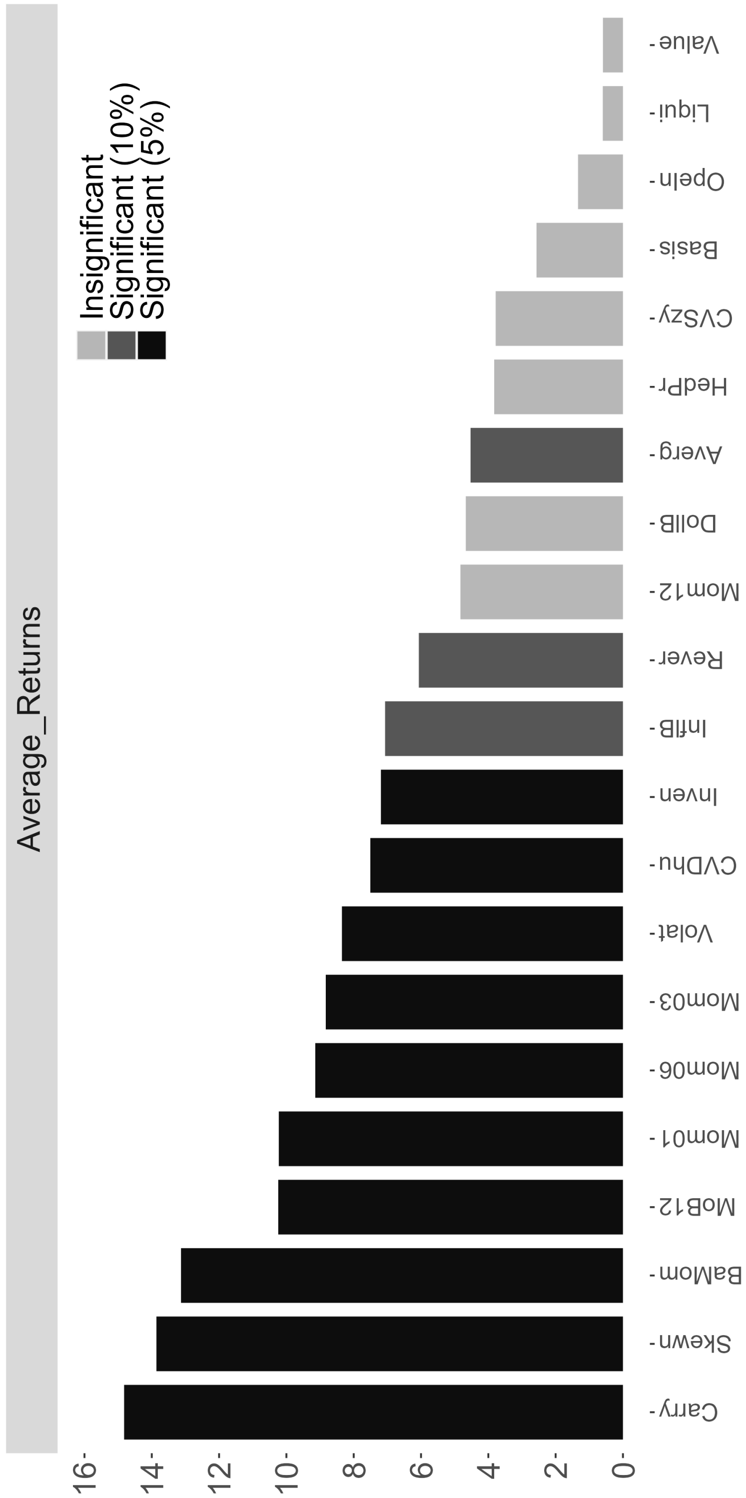


Figure A.2: Average returns to the commodity investment strategies (Full Sample)
This histogram shows the average returns to each of the commodity trading strategies. The returns are annualized excess returns in percentage points. The sample period covers monthly data from March 1986 to August 2021. 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) is analyzed in Table 1.

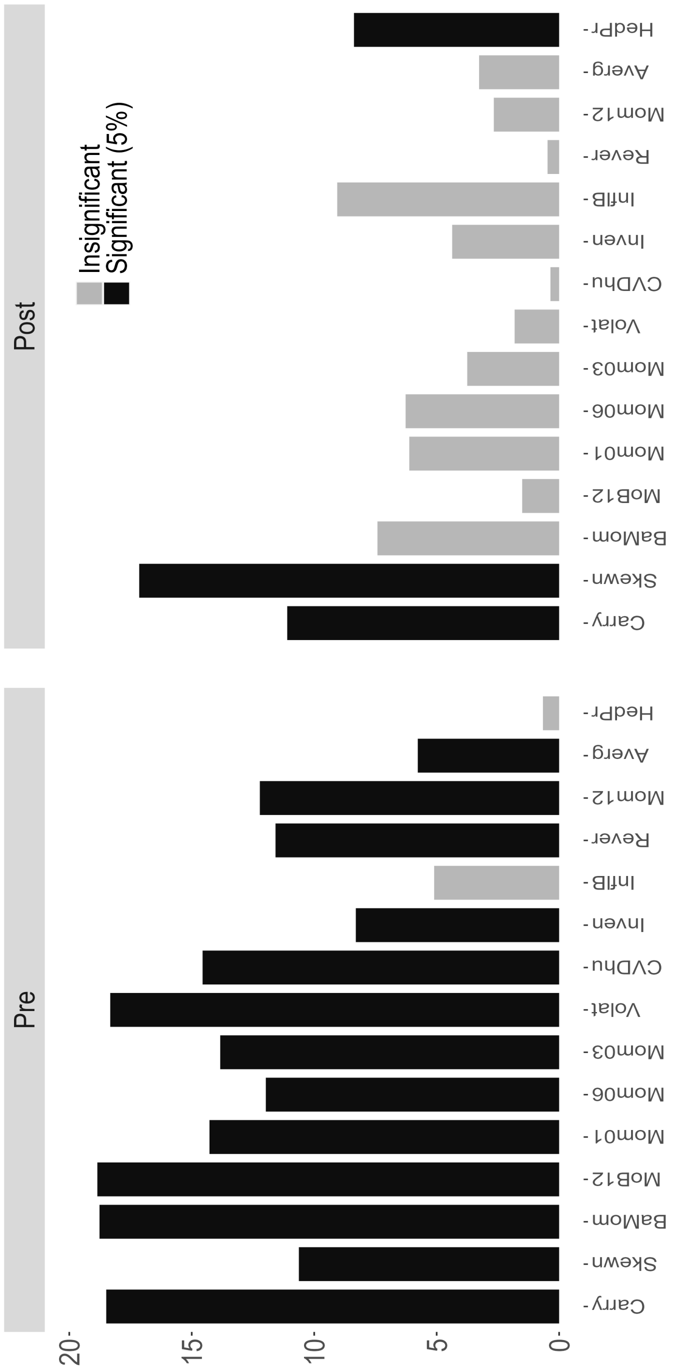


Figure A.3: Average returns to the commodity investment strategies (Pre and Post-Financialization)

This histogram shows the average returns to each of the commodity trading strategies. The returns are annualized excess returns in percentage points. The sample period covers monthly data 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) is analyzed in Table 1.

A.5 Date of Financialization

As mentioned in the main part of the paper, the consensus in the literature is to date the financialization to the beginning of 2004 (see Basak and Pavlova (2016), Brogaard et al. (2019), Goldstein and Yang (2021), among others). This choice is typically justified by the observed spike in commodity investments around that period (see figure Figure A.6).

To further substantiate the chosen date of financialization in this study, we conduct a hierarchical Bayesian analysis using the returns of the commodity trading strategies. This methodology allows us to infer the most likely date of financialization by examining changes in the average returns of the strategies.

We begin by specifying an uninformative prior distribution for the date of financialization. Given our sample period, we assume each date is equally likely and define the breakpoint (date of financialization) τ as follows:

$$P(\tau) \sim \text{Uniform}[1986/03/31,\ 2021/07/30], \tag{20}$$

Let $r_{i,t}$ denote the return of a commodity trading strategy i at time t. We assume that each return process follows a strategy-specific normal distribution:

$$r_{i,t} \sim \mathcal{N}(\mu_{\text{i}}, \sigma_i^2) \tag{21}$$

where μ_{i} and σ_i^2 are the mean and variance of the returns of strategy i.

Our interest is in testing whether or not the average returns of commodity trading strategies change around an unknown date τ. To formalize this, we allow for a possible change in strategy-specific average returns before and after the breakpoint:

$$r_{i,t} \sim \begin{cases} \mathcal{N}(\mu_{\text{i, pre}}, 1.0), & for t \leq \tau, \\ \mathcal{N}(\mu_{\text{i, post}}, 1.0), & for t > \tau, \end{cases} \tag{22}$$

We also assign a Half-Cauchy prior on the standard deviation of the returns for all strategies before and after the breakpoint:

$$\sigma_i \sim \text{Half-Cauchy}(0, 5.0) \tag{23}$$

Our primary interest lies in the posterior distribution of the breakpoint τ, which represents our updated beliefs about the timing of financialization's impact on commodity futures strategies, given the observed returns. We obtain this distribution by applying Bayes' rule as follows:

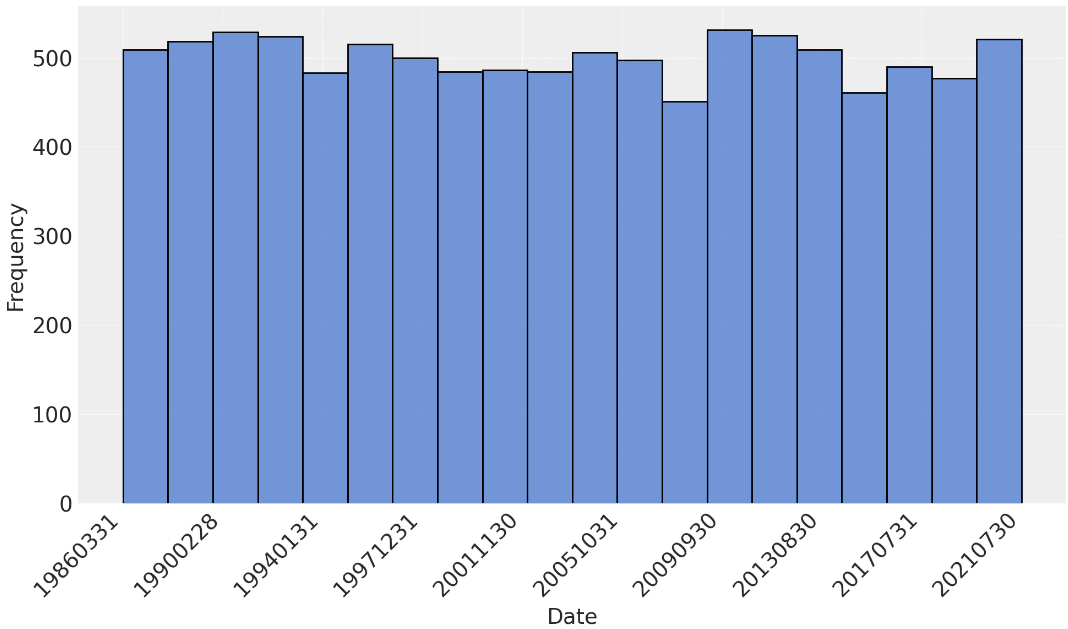
$$P(\tau|\text{data}) = \frac{P(\text{data}|\tau) \cdot P(\tau)}{P(\text{data})}, \tag{24}$$

where $P(\text{data}|\tau)$ is the likelihood of observing the data given the breakpoint τ and distributional assumptions on the data, and $P(\text{data})$ serves as a normalization constant. To compute the likelihood of the data given the breakpoint, we integrate out the strategy returns:

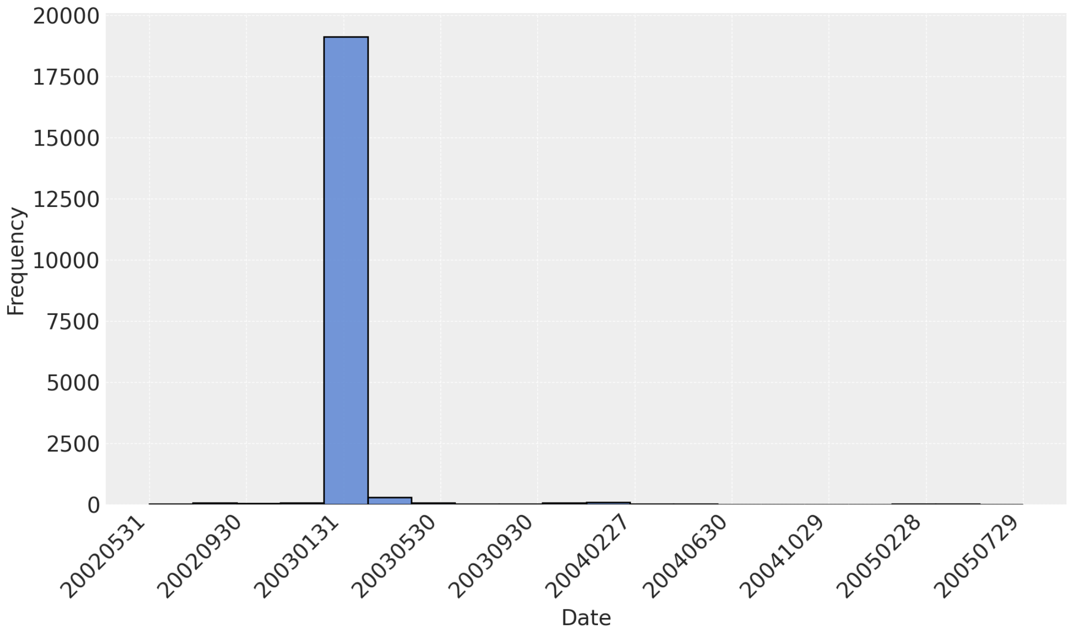
$$P(\text{data}|\tau) = \int P(\text{data}|\mu_{\text{pre}}, \mu_{\text{post}}, \sigma) \cdot P(\mu_{\text{pre}}, \mu_{\text{post}}, \sigma|\tau)\, d\mu_{\text{pre}}\, d\mu_{\text{post}}\, d\sigma, \tag{25}$$

We implement this integral using a Markov Chain Monte Carlo (MCMC) algorithm, specifically the No-U-Turn Sampler (NUTS), to sample from the posterior distribution of the breakpoint τ.

Figure A.4 shows the posterior distribution of the breakpoint τ, which is centered around January 2003. About 95% of the posterior mass sits on January 2003. This is about 11 months ahead of the current consensus date in the literature, and so we choose to stick with the literature date of 2004. This should make our results more comparable with existing studies. Moreover, a Wald test (from the frequentist perspective) to determine if the returns of the trading strategies are jointly different before and after 2004 would also confirm that the returns statistically change around the financialization date.



(a) Prior Distribution of Financialization Date



(b) Posterior Distribution of Financialization Date

Figure A.4: Distribution of Financialization Date

Figure 2 presents the results of our Bayesian analysis to infer the most likely date of financialization based on changes in the average returns of commodity trading strategies. Panel (a) shows the uninformative prior distribution, which assumes each date between March 31, 1986, and July 30, 2021, is equally likely to be the date of financialization. Panel (b) depicts the posterior distribution of the financialization date after updating our beliefs using the observed returns data. The posterior distribution is centered around 2003. This finding supports our choice of using the literature date of 2004 as the breakpoint for the pre- and post-financialization periods in our main analysis.

A.6 Spanning Tests

Table A.2: Spanning Tests (T-Statistics of the Intercepts)

This table displays the t-statistics of the intercepts resulting from regressions of the returns to each strategy on a constant and another strategy. We compute test statistics (*tstat*) using Newey and West (1987) corrected standard errors (with lag selection following Andrews (1991)). 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. The data contain monthly series from March 1986 to August 2021.

	Averg	Carry	Mom06	Mom12	InflB	DollB	CVSzy	CVDhu	HedPr	BaMom	MoB12	Basis	Volat	Value	Inven	Skewn	OpeIn	Liqui	Mom03	Mom01	Rever
Averg	0.00	1.13	1.72	1.64	1.30	1.49	1.89	1.48	1.66	1.52	1.60	1.82	0.93	1.89	2.27	1.31	1.95	1.83	1.74	1.81	1.95
Carry	1.13	0.00	4.36	4.76	4.82	4.79	4.55	4.50	4.79	3.80	4.34	5.71	4.88	5.22	4.19	3.75	4.87	4.96	4.26	4.46	4.92
Mom06	1.72	4.36	0.00	1.82	2.32	2.49	2.21	2.00	2.23	1.84	0.96	2.76	2.33	2.87	2.73	2.46	2.57	2.50	1.09	1.25	2.23
Mom12	1.64	4.76	1.82	0.00	1.06	1.45	1.23	0.67	1.21	0.38	-0.73	1.67	1.07	1.76	2.53	1.01	1.40	1.40	0.47	0.87	1.16
InflB	1.30	4.82	2.32	1.06	0.00	1.48	1.84	1.79	1.80	1.48	1.75	1.93	1.52	2.01	1.44	1.22	1.89	1.93	1.70	2.01	2.37
DollB	1.49	4.79	2.49	1.45	1.48	0.00	-1.38	-1.34	-1.24	-1.01	-1.57	-1.54	-1.42	-1.43	-0.55	-0.81	-1.53	-1.45	-1.48	-1.59	-1.78
CVSzy	1.89	4.55	2.21	1.23	1.84	-1.38	0.00	1.03	1.06	0.81	0.82	1.36	1.42	1.27	1.35	1.97	1.23	1.24	0.30	0.61	1.20
CVDhu	1.48	4.50	2.00	0.67	1.79	-1.34	1.03	0.00	2.44	2.39	2.12	2.61	2.00	2.63	2.26	1.94	2.54	2.53	2.18	2.22	2.60
HedPr	1.66	4.79	2.23	1.21	1.80	-1.24	1.06	2.44	0.00	1.37	1.19	1.37	1.23	1.29	0.55	0.60	1.17	1.25	0.97	1.10	1.39
BaMom	1.52	3.80	1.84	0.38	1.48	-1.01	0.81	2.39	1.37	0.00	2.37	3.31	2.87	2.86	3.30	2.77	2.93	2.96	2.54	2.68	2.57
MoB12	1.60	4.34	0.96	-0.73	1.75	-1.57	0.82	2.12	1.19	2.37	0.00	2.48	2.10	2.44	3.24	1.99	2.21	2.18	1.07	1.20	1.85
Basis	1.82	5.71	2.76	1.67	1.93	-1.54	1.36	2.61	1.37	3.31	2.48	0.00	0.50	0.53	1.02	0.29	0.57	0.58	1.34	1.18	0.33
Volat	0.93	4.88	2.33	1.07	1.52	-1.42	1.42	2.00	1.23	2.87	2.10	0.50	0.00	2.37	2.72	1.65	2.37	2.33	2.15	2.30	2.38
Value	1.89	5.22	2.87	1.76	2.01	-1.43	1.27	2.63	1.29	2.86	2.44	0.53	2.37	0.00	-0.02	0.51	0.16	0.17	1.07	0.76	-0.15
Inven	2.27	4.19	2.73	2.53	1.44	-0.55	1.35	2.26	0.55	3.30	3.24	1.02	2.72	-0.02	0.00	-2.87	-2.57	-2.65	-2.78	-2.77	-2.89
Skewn	1.31	3.75	2.46	1.01	1.22	-0.81	1.97	1.94	0.60	2.77	1.99	0.29	1.65	0.51	-2.87	0.00	3.91	4.02	4.24	4.39	4.34
OpeIn	1.95	4.87	2.57	1.40	1.89	-1.53	1.23	2.54	1.17	2.93	2.21	0.57	2.37	0.16	-2.57	3.91	0.00	-0.38	-0.86	-0.73	-0.39
Liqui	1.83	4.96	2.50	1.40	1.93	-1.45	1.24	2.53	1.25	2.96	2.18	0.58	2.33	0.17	-2.65	4.02	-0.38	0.00	-0.58	-0.33	-0.02
Mom03	1.74	4.26	1.09	0.47	1.70	-1.48	0.30	2.18	0.97	2.54	1.07	1.34	2.15	1.07	-2.78	4.24	-0.86	-0.58	0.00	0.85	1.95
Mom01	1.81	4.46	1.25	0.87	2.01	-1.59	0.61	2.22	1.10	2.68	1.20	1.18	2.30	0.76	-2.77	4.39	-0.73	-0.33	0.85	0.00	2.14
Rever	1.95	4.92	2.23	1.16	2.37	-1.78	1.20	2.60	1.39	2.57	1.85	0.33	2.38	-0.15	-2.89	4.34	-0.39	-0.02	1.95	2.14	0.00

A.7 Styled Fact - Robustness (Date Before Financialization of Commodity Markets)

Table A.3: 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 the 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.8 Equity Trading Strategies

In this section, we explore strategies in commodity futures markets that are inspired by the equity literature but have not been previously implemented in the existing commodity literature. Specifically, we adapt strategies outlined in Chen and Zimmermann (2020) for application to commodities whenever feasible. We retain the following set of strategies, for which we refer to their paper for further details on construction:

- 52 week high;
- CAPM beta;
- Downside beta;
- Idiosyncratic risk;
- Intermediate momentum
- Kurtosis;
- Lottery;
- Maximum return over month;
- Momentum-Reversal;
- Momentum and Volume
- Past trading volume;
- Price;
- Seasonality;
- Tail risk beta;
- Volume trend;
- Volume variance;

Overall, we observe a decline in the profitability of these trading strategies around financialization (see Figure Table A.5 below). However, the dynamics of a couple of strategies tend to now move in the opposite direction to our baseline findings, i.e., they gain profitability after financialization. One of these strategies is CAPM beta, and it's well-established that the correlation between commodities and equities increased around financialization (see, for example Tang and Xiong (2012) and Boons et al. (2014)).

Moreover, in untabulated results, we regress the returns to these strategies on a dummy variable that takes the value one after financialization and zero before. We find that the strategies tend to experience a loss in profitability of around 20 basis points per month around financialization. Although this result is not statistically significant and quantitatively half of what we find for the baseline commodity strategies in Table 4, it qualitatively reinforces the evidence of a decay in excess returns in commodity trading strategies around financialization.

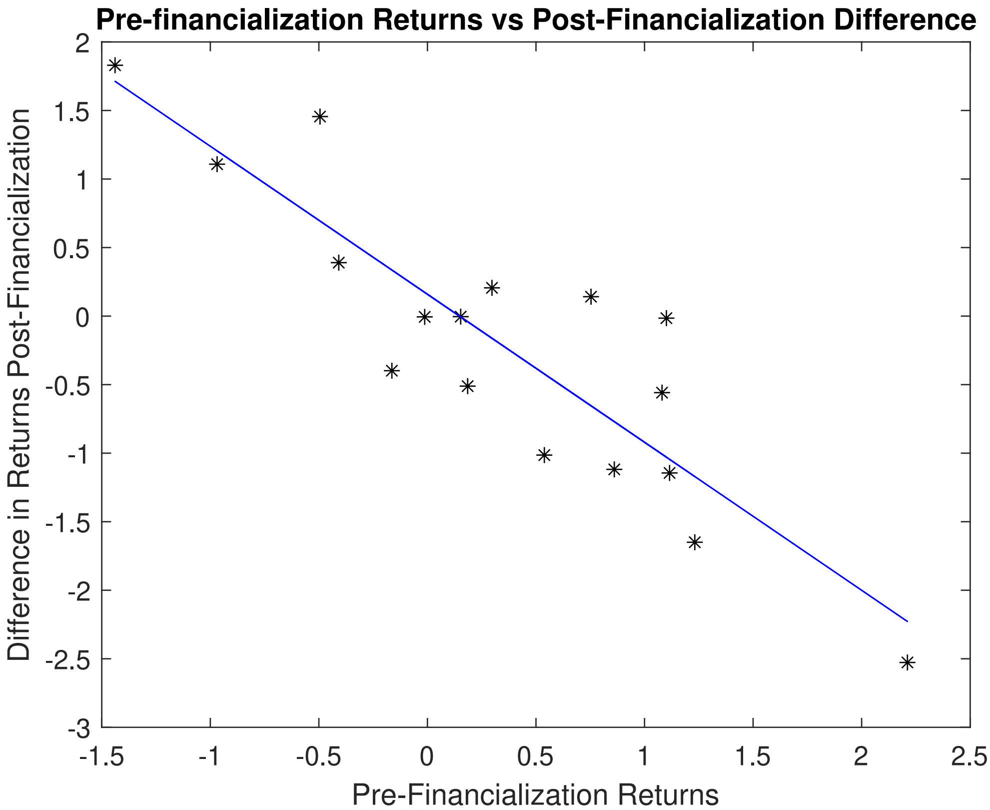


Figure A.5: Relation between pre- and post-financializtion returns to equity strategies
This scatter plot shows the relation between the monthly returns to the 17 equity strategies (we apply to commodities) pre-financialization and the changes in their returns post-financialization. The returns pre-financialization 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 period covers monthly data from March 1986 to August 2021.

A.9 Unconditional Asset Pricing Test - 1st Stage Regressions

Table A.4: 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. Specifically, we report ordinary least squares (OLS) estimates of contemporaneous time-series regressions of the strategies on the latent risk factors (along with 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 (i.e., 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 are expressed in percentage per month. 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. 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.10 Total Open Interest in Commodity Markets

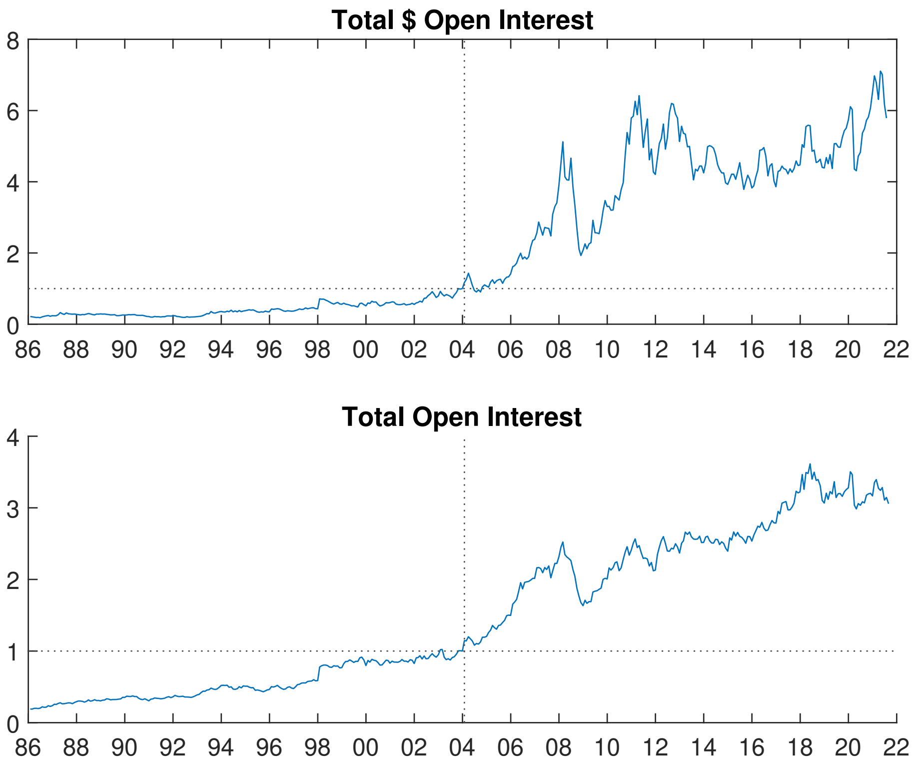


Figure A.6: Total ($) Open Interest

This figure plots the sum of dollar open interest and open interest over time for the commodities in our sample. The data cover the period March 1986 to August 2021 and are retrieved from the CFTC. Both series are normalized to equal one in December 2003, i.e., at the outset of the financialization.

A.11 Unconditional Asset Pricing Test - Higher Number Of Latent Factors

Table A.5: Unconditional Asset Pricing Tests (Over Subsamples) - 8 Latent-Factor Model

This table presents the results for the second (cross-sectional) stage of the Fama and MacBeth (1973) asset pricing tests, employing a higher number of latent factors compared to Table 2. We use as test assets the returns to the commodity investment strategies presented in Table 1; while, the eight candidate factors are the first eight RP-PCs extracted as in Lettau and Pelger (2020). *Panel A* reports the results for the test conducted over the full sample (i.e.,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. We compute test statistics (*tstat*) using Newey and West (1987) corrected standard errors (with lag selection following Andrews (1991)). 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.12 Characterising the latent factors

We here characterize the properties of the latent factor model presented in Section 3 which, as shown, thorugh a handful of systematic factors is able to account for the dynamics of the returns to the commodity strategies. 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 mainly focus our analysis on the properties of these latent factors.

A.12.1 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 A.6.

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. 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. 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.

Table A.6: 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. We compute test statistics using Newey and West (1987) corrected standard errors (with lag selection following 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

A.12.2 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:

$$RPPC_t^j = \phi_0 + \phi_1 \Delta MacroFin_t^i + \epsilon_t \tag{26}$$

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.

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[21]. We present the results from this exercise in Table A.8.

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[22]. 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

[21] Table A.7 lists and describes more in details the variables and the sources from which they are retrieved.

[22] 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).

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 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.9 and Table A.10 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.

A.12.3 Macro-Financial Variables - Tables

Table A.7: 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 A.8. The second column (*Category*) reports to which broad category the variables belong to. The third column (*Description*) provides concise descriptions of the variables along with 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.8: 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. We compute test statistics 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. Table A.7 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 A.9: 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). We compute test statistics 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. Table A.7 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.10: 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). We compute test statistics 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. Table A.7 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%

Table A.11: Descriptive Statistics - Returns to the Commodity Investment Strategies (Over Subsamples, Excluding GFC and Covid)

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. However, it's important to note that in this table, the analysis conducted over the full sample and post-financialization period (i.e., the first and third panels) excludes the Global Financial Crisis and Covid periods. 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 the 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***	14.89	18.71	0.80	4.69	Carry***	18.49	20.64	0.90	3.95	Carry**	10.17	15.78	0.64	2.40
BaMom***	17.18	23.99	0.72	4.33	BaMom***	18.77	25.77	0.73	3.17	BaMom***	15.09	21.48	0.70	2.88
Skewn***	12.62	20.65	0.61	3.47	Skewn*	10.63	22.89	0.46	1.89	Skewn***	15.24	17.31	0.88	3.57
CVDhu***	9.25	17.01	0.54	3.14	CVDhu***	14.56	17.69	0.82	3.76	CVDhu	2.29	15.89	0.14	0.56
Volat***	11.09	23.12	0.48	3.13	Volat***	18.33	25.03	0.73	3.89	Volat	1.54	20.08	0.08	0.30
MoB12**	11.43	28.38	0.40	2.30	MoB12***	18.85	31.26	0.60	2.63	MoB12	1.69	23.89	0.07	0.28
Inven**	-7.28	14.27	-0.51	-2.30	Inven**	-8.30	14.26	-0.58	-2.36	Inven	-2.73	14.37	-0.19	-0.38
Averg**	4.56	10.31	0.44	2.24	Averg***	5.77	9.47	0.61	2.58	Averg	2.97	11.32	0.26	0.91
Mom06*	7.66	24.11	0.32	1.95	Mom06**	11.98	25.93	0.46	2.05	Mom06	1.98	21.45	0.09	0.45
Mom12*	6.72	21.23	0.32	1.88	Mom12**	12.22	22.92	0.53	2.40	Mom12	-0.49	18.65	-0.03	-0.11
Mom03*	8.31	24.29	0.34	1.80	Mom03**	13.84	25.27	0.55	2.25	Mom03	1.06	22.85	0.05	0.17
Mom01*	8.26	24.42	0.34	1.77	Mom01**	14.28	26.28	0.54	2.34	Mom01	0.36	21.60	0.02	0.06
Rever	6.43	22.28	0.29	1.63	Rever**	11.58	23.02	0.50	2.25	Rever	-0.33	21.19	-0.02	-0.05
InflB	6.09	20.97	0.29	1.61	InflB	5.10	19.98	0.26	1.09	InflB	7.38	22.26	0.33	1.27
CVSzy	3.43	17.57	0.20	1.06	CVSzy	5.23	19.33	0.27	1.08	CVSzy	1.08	14.97	0.07	0.29
Basis	3.41	25.70	0.13	0.75	Basis	5.40	27.94	0.19	0.78	Basis	0.78	22.48	0.03	0.14
DollB	-2.54	19.06	-0.13	-0.75	DollB	-2.29	18.81	-0.12	-0.50	DollB	-2.86	19.43	-0.15	-0.58
HedPr	1.97	17.60	0.11	0.66	HedPr	-0.66	18.24	-0.04	-0.15	HedPr	5.43	16.72	0.32	1.59
Value	1.82	17.61	0.10	0.53	Value	-0.05	18.40	0.00	-0.01	Value	4.28	16.55	0.26	0.94
OpeIn	-1.71	18.68	-0.09	-0.52	OpeIn	0.19	21.03	0.01	0.04	OpeIn	-4.21	15.09	-0.28	-1.24
Liqui	-0.46	15.65	-0.03	-0.16	Liqui	-0.15	16.02	-0.01	-0.04	Liqui	-0.86	15.22	-0.06	-0.23

A.13 Returns

The geometric drift and diffusion terms for the returns process are obtained from applying Ito's lemma to the definition of returns. We define returns as:

$$\begin{aligned} dR_{j,t} &= \frac{d(p_{j,t}Y_{j,t}/F_{j,t})}{(p_{j,t}Y_{j,t}/F_{j,t})} + F_{j,t}dt \\ dR_{j,t} &= \mu_{R_j,t}dt + \sigma'_{R_j,t}d\vec{Z}_t \end{aligned} \tag{27}$$

using remark ??. Define generic Ito processes for $p_{j,t}$. as:

$$dp_{j,t} = p_{j,t}\left(\mu_{p_j,t}dt + \sigma_{p_j,t}d\vec{Z}_t\right) \tag{28}$$

Using the results for the product of two Ito processes, we have:

$$\begin{aligned} \frac{d\left((p_{j,t}Y_{j,t})\right)}{(p_{j,t}Y_{j,t})} &= \left(\mu_{p_j,t} + \mu_{Y_j} + \sigma_{p_j,t}\sigma_{Y_j}\right)dt + \left(\sigma_{p_j,t} + \sigma_{Y_j}\right)d\vec{Z}_t \\ \frac{d\left((p_{j,t}Y_{j,t})\right)}{(p_{j,t}Y_{j,t})} &= \mu_{py,t}dt + \sigma_{py,t}d\vec{Z}_t \\ \frac{d\left(py_{j,t}\right)}{(py_{j,t})} &= \mu_{py_j,t}dt + \sigma_{py_j,t}d\vec{Z}_t \end{aligned} \tag{29}$$

where we have simply redefined $py_{j,t} = \frac{d((p_{j,t}Y_{j,t}))}{(p_{j,t}Y_{j,t})}$ for notational convenience. Again applying Ito's Lemma to $d\left(py_{j,t}\right)/F_{j,t}$, we have:

$$\frac{d\left(py_{j,t}\right)/F_{j,t}}{\left(py_{j,t}\right)/F_{j,t}} = \left(\mu_{py,t} - \mu_{F_j,t} - \sigma'_{py,t}\sigma_{f,t}\right)dt + \left(\sigma_{py,t} - \sigma_{F_j,t}\right)'d\vec{Z}_t \tag{30}$$

Substituting drift and volatility terms from 29 into 30 and then into 27, we have:

$$\begin{aligned} dR_{i,t} = &\left(F_{j,t} + \mu_{p_j,t} + \mu_{Y_j} + \sigma'_{p_j,t}\sigma_{Y_j} + \sigma_{F_j,t}{}'\sigma_{F_j,t} - \mu_{F_j,t} - \left(\sigma_{p_j,t} + \sigma_{Y_j}\right)'\sigma_{F_j,t}\right)dt \\ &+ (\sigma_{p_j,t} + \sigma_{Y_j} - \sigma_{F_j,t})'d\vec{Z}_t \end{aligned} \tag{31}$$

Index Returns

We define the index as a value-weighted portfolio of Lucas trees. Let us focus on the case of two Lucas trees. Let $Q_{1,t}$ and $Q_{2,t}$ be the prices of the two trees. The total index value is:

$$Q_t = Q_{1,t} + Q_{2,t}.$$

Because each good j is sold at price $p_{j,t}$, the combined dividend paid out by the index in terms of the numeraire is:

$$p_{1,t}Y_{1,t} + p_{2,t}Y_{2,t}.$$

The total index return dR_t is defined analogously to the single-tree case, as:

$$dR_t = \frac{dQ_t}{Q_t} + \frac{p_{1,t}Y_{1,t} + p_{2,t}Y_{2,t}}{Q_t}dt.$$

The index return consists of the capital gain (dQ_t/Q_t) plus the dividend yield from both assets aggregated together, $(p_{1,t}Y_{1,t} + p_{2,t}Y_{2,t})/Q_t$.

We define the portfolio weights of the index in terms of the value shares of the individual assets:

$$s_t = \frac{Q_{1,t}}{Q_t}, \quad 1 - s_t = \frac{Q_{2,t}}{Q_t}.$$

With these weights, we can express the index return as a weighted combination of the returns of the two individual assets. Let the return on tree j be $dR_{j,t}$:

$$dR_{j,t} = \frac{dQ_{j,t}}{Q_{j,t}} + \frac{p_{j,t}Y_{j,t}}{Q_{j,t}}dt.$$

By construction, the index return is just the value-weighted average of the two tree returns:

$$dR_t = s_t dR_{1,t} + (1 - s_t)dR_{2,t}.$$

This relationship is key—it allows us to use the results obtained for individual trees to determine the index's return process.

From the single-tree derivation (see Appendix A.13), we know that each tree's return can be written in the Ito form as:

$$dR_{j,t} = \mu_{R_j,t}dt + \sigma'_{R_j,t}d\vec{Z}_t,$$

where $\mu_{R_j,t}$ and $\sigma_{R_j,t}$ are the drift and diffusion terms for the j-th asset's return. Each $\mu_{R_j,t}$ encapsulates the effect of the dividend yield, the price dynamics of the underlying asset, and the Ito correction terms. Similarly, $\sigma_{R_j,t}$ gives the response of the return to the Brownian shocks $d\vec{Z}_t$.

Substitute these into the index return:

$$dR_t = s_t(\mu_{R_1,t}dt + \sigma'_{R_1,t}d\vec{Z}_t) \ + \ (1 - s_t)(\mu_{R_2,t}dt + \sigma'_{R_2,t}d\vec{Z}_t).$$

Grouping terms, we obtain:

$$dR_t = [s_t\mu_{R_1,t} + (1 - s_t)\mu_{R_2,t}]dt + [s_t\sigma_{R_1,t} + (1 - s_t)\sigma_{R_2,t}]'d\vec{Z}_t.$$

From the above, we see that the drift of the index return, $\mu_{R,t}$, is a weighted average of the individual assets' return drifts:

$$\mu_{R,t} = s_t\mu_{R_1,t} + (1 - s_t)\mu_{R_2,t}.$$

Similarly, the diffusion (volatility) term is also a weighted average:

$$\sigma_{R,t} = s_t\sigma_{R_1,t} + (1 - s_t)\sigma_{R_2,t}.$$

Because the index is value-weighted, its drift and volatility naturally aggregate the underlying assets' return characteristics proportionally to their share in the index.

We can rewrite the dividend yield portion of the index return more explicitly. Define the index dividend yield F_t as:

$$F_t = \frac{p_{1,t}Y_{1,t} + p_{2,t}Y_{2,t}}{Q_t}.$$

Since $Y_t = Y_{1,t} + Y_{2,t}$ and $Q_t = Q_{1,t} + Q_{2,t}$, the index's dividend yield is also a value-weighted combination of the individual dividend yields $F_{j,t} = \frac{p_{j,t}Y_{j,t}}{Q_{j,t}}$:

$$F_t = s_t F_{1,t} + (1 - s_t)F_{2,t}.$$

This ties the interpretation neatly back to the fundamentals: the index is a portfolio with weights s_t and $(1 - s_t)$, and hence its return components—both drift and dividend yield—inherit this weighting structure.

The return process for the index can be expressed as:

$$dR_t = \mu_{R,t}dt + \sigma'_{R,t}d\vec{Z}_t,$$

with

$$\mu_{R,t} = s_t\mu_{R_1,t} + (1 - s_t)\mu_{R_2,t}, \quad \sigma_{R,t} = s_t\sigma_{R_1,t} + (1 - s_t)\sigma_{R_2,t}.$$

Since $s_t = \frac{Q_{1,t}}{Q_t}$, the drift and volatility of the index return are determined by the individual drifts and volatilities of the component assets, weighted by their relative importance in the total index value.

A.14 State variables

Wealth share

The first variable that summarises the state of the economy is the wealth share of the first investor. The wealth share is a relevant state-variable in our economy because wealth is not constant as in standard log utility models. Additionally, wealth is not a monotonic function of other fundamentals in the economy. Wealth share fundamentally captures time-varying Negishi weights as discussed in Dumas et. 2000, Anderson (2005) or Colcaito and Croce (2011, 2013).

$$
\begin{aligned}
x_{i,t} &= \frac{W_t^i}{\sum_{k=0}^{M-1} W_t^k} \\
x_{i,t} &= \frac{x_{i,t}^1}{x_t^2}
\end{aligned}
$$

We solve for dynamics by defining the numerator and denomitor separately and solve for them.

$$
x_t^2 = \sum_{k=0}^{M-1} W_t^k \tag{32}
$$

$$
dx_t^2 = \sum_{i=0}^{M-1} \mu_i W_t^i dt + \sum_{i=0}^{M-1} W_t^i \sigma_i' d\vec{Z}_t \tag{33}
$$

we denote the drift and volatility of the denominator as:

$$
\begin{aligned}
\mu_{x_t^2,t} &= \sum_{j=0}^{M-1} \frac{\mu_j W_t^i}{x_t^2} \\
\sigma_{x_t^2,t} &= \sum_{i=0}^{M-1} \frac{W_t^i}{x_t^2} \sigma_j
\end{aligned}
$$

The numerator is straightforward:

$$
x_{j,t}^1 = W_t^i \tag{34}
$$

$$
dx_t^1 = \mu_{w_i,t} W_t^i dt + W_t^i \sigma_{w_i,t}' d\vec{Z}_t \tag{35}
$$

we denote the drift and volatility of the numerator as:

$$\mu_{x^1_t,t} = \mu_{w_i,t}$$
$$\sigma_{x^1_t,t} = \sigma_{w_i,t}$$

Then the dynamics of $x_{i,t}$ follows:

$$\frac{dx_{j,t}}{x_{j,t}} = \left(\mu_{w_i,t} - \mu_{x^2_t,t} + \sigma'_{x^2_t,t}(\sigma_{x^2_t,t} - \sigma_{w_i,t})\right) dt + \left(\sigma_{w_i,t} - \sigma_{x^2_t,t}\right)' d\vec{Z}_t \tag{36}$$
$$\frac{dx_{j,t}}{x_{j,t}} = \mu_{x,t} x_t dt + x_t \sigma'_{x,t} d\vec{Z}_t$$

Realtive supply of fruits

The second state variable is the relative supply of good 1. This is defined as:

$$y_t = \frac{Y_{1,t}}{Y_{1,t} + Y_{2,t}} \tag{37}$$

We $W^i_t \geq 0$ and $Y_{j,t} \geq 0$ for all t and i, j and so x_t and y_t evolve in the bounded interval $[0, 1]$.

$$y_{j,t} = \frac{Y_{j,t}}{\sum_{k=0}^{M-1} Y_{k,t}}$$
$$y_{j,t} = \frac{y^1_{j,t}}{y^2_t}$$

We solve for dynamics by defining the numerator and denomitor separately and solve for them.

$$y^2_t = \sum_{j=0}^{M-1} Y_{j,t} \tag{38}$$

$$dy_t^2 = \sum_{j=0}^{M-1} \mu_j Y_{j,t} dt + \sum_{j=0}^{M-1} Y_{j,t} \sigma_j' d\vec{Z}_t \tag{39}$$

we denote the drift and volatility of the denominator as:

$$\mu_{y_t^2,t} = \sum_{j=0}^{M-1} \frac{\mu_j Y_{j,t}}{y_t^2}$$
$$\sigma_{y_t^2,t} = \sum_{j=0}^{M-1} \frac{Y_{j,t}}{y_t^2} \sigma_j$$

The numerator is straightforward:

$$y_{j,t}^1 = Y_{j,t} \tag{40}$$

$$dy_{j,t}^1 = \mu_j Y_{j,t} dt + Y_{j,t} \sigma_j' d\vec{Z}_t \tag{41}$$

we denote the drift and volatility of the numerator as:

$$\mu_{y_t^1,t} = \mu_{Y_j}$$
$$\sigma_{y_t^1,t} = \sigma_{Y_j}$$

Then the dynamics of $y_{j,t}$ follows:

$$\frac{dy_{j,t}}{y_{j,t}} = \left(\mu_j - \mu_{y_t^2,t} + \sigma_{y_t^2,t}'(\sigma_{y_t^2,t} - \sigma_{Y_j})\right) dt + \left(\sigma_{Y_j} - \sigma_{y_t^2,t}\right)' d\vec{Z}_t \tag{42}$$
$$\frac{dy_{j,t}}{y_{j,t}} = \mu_{y,t} y_t dt + y_t \sigma_{y,t}' d\vec{Z}_t$$