

Media emotion intensity and commodity futures pricing

Abstract

This study investigates the impact of media emotion intensity on commodities futures returns. Emotion intensity measures the proportion of emotional content relative to factual content in media news. This new factor is distinct from news sentiment and media coverage and exhibits weak cross-sectional correlations with both of them. The media emotion intensity factor exhibits an annual premium of around 14.40% and is more pronounced for commodities with low media coverage, high momentum, high basis-momentum, high hedging pressure, and backwardation. Emotion intensity significantly predicts the trading tendencies of both commercial and non-commercial traders and the cross-section of commodity futures returns at both portfolio and individual levels. Further, other commonly considered risk sources cannot subsume the predictability of the media emotion intensity factor.

Keywords: commodity futures, media emotion intensity, asset pricing

JEL: G110, G130, G40

1. Introduction

In recent decades, commodity futures have become increasingly important in investor portfolios, with a growing trend towards financialization (Adams and Glück, 2015; Belousova and Dorfleitner, 2012; Tang and Xiong, 2012). This explains the increase in research on exploring the factors that can be used as signals for investment and can explain the cross-sectional returns of commodity futures. Key benchmarks, including basis (Szymanowska et al., 2014), momentum (Bakshi et al., 2019), basis-momentum (Boons and Prado, 2019), and other factors, have been introduced. While equity market research has proposed over 300 signals, the number of suggested factors in the commodity market that can predict the cross-section of returns remains limited.

This study examines the impact of media emotion intensity on the cross-section of commodity futures return. Existing literature on the role of media in asset pricing has documented the effects of two media factors: news sentiment (García, 2013; Smales, 2014; Tetlock, 2007) and media coverage (Fang and Peress, 2009). News sentiment is built on the proportion of positive words and phrases to their negative counterparts. This media news sentiment explains asset price movement in two ways. On the one hand, positive and negative levels of media news may be associated with positive and negative information in news content that fundamentally explains asset prices (Tetlock et al., 2008). On the other hand, news sentiment can trigger investors'

behavioral biases, which drive price movement (García, 2013; Tetlock, 2007). Meanwhile, media coverage measures the volume of news articles that mention a given asset. Fang and Peress (2009) investigate this factor and suggest that stocks without news in the media generate significantly higher returns than those highly featured in the media.

This research focuses on another aspect of media, media emotion, deviating from frequently examined factors such as news sentiment and media coverage. Recent academic research has also highlighted the growing importance of using emotion in journalism, particularly in the digital and social media era (Wahl-Jorgensen, 2020). The media strategically employs emotion to attract attention and engage readers (Beckett, 2015). Goldenberg and Gross (2020) have demonstrated that media can transmit emotions to readers, and increased exposure to emotional content can make individuals more emotional. This suggests that exposure to a higher intensity of emotion in media news can have a significant impact on reader judgments and decisions. Neuroscientific research has shown that humans tend to prioritize emotional stimuli, making it difficult to ignore emotion in decision-making processes (Vuilleumier, 2005). While the literature is inconclusive on whether emotion impairs or enhances judgment (Lerner et al., 2015; Pham, 2007), studies indicate that emotion can trigger behavioral biases in judgment (Han et al., 2007; Loewenstein, 1996; Loewenstein and Lerner, 2003). Stronger emotions associated with decision-making can override rational behaviors (Loewenstein et al., 2001). Meanwhile, some studies suggest that emotion can enhance judg-

ment (Schwarz, 2000; Solomon, 1993). In finance and economics, emotional factors are necessary in interpreting economic and risk-taking behaviors (Da et al., 2015; Griffith et al., 2020; Kuhnen and Knutson, 2011; Loewenstein, 2000). However, these current studies primarily concentrate on specific emotional aspects, documenting their influence on financial decision-making and asset pricing. Our research further focuses on another unique facet of media emotion, namely media emotion intensity, which reflects the ratio of emotional to factual content in news items, diverging from particular emotional aspects. Even when the media shows more fear or joy towards an asset at a certain time, it does not necessarily mean that the proportion of emotional content to factual content in the media news is higher compared to times when the media news shows less fear and joy.

In this research, our main theme is the impact of media emotion intensity on commodity futures returns. To construct media emotion intensity for each commodity, the *EmotionVsFact* index from the Thomson Reuters *MarketPsych* data set is employed. The index calculates the excess proportion of emotional content over factual content, scaled by the total amount of content and adjusted for the total amount of media coverage. Because a higher level of specific emotional aspects in news items is not necessarily associated with a higher ratio of emotional to factual content, media emotion intensity and specific emotional aspects should be distinguished. The media emotion intensity factor aims to capture the emotional intensity of news articles, regardless of whether the emotional words and phrases used

are positive or negative. The low average time-series correlation of 0.099 and cross-section correlation of 0.160 between media emotion intensity and news sentiment in the research sample imply that the emotion intensity factor we present in this paper is clearly different from the sentiment factor in prior literature.

First, this study examines the media emotion intensity's impact on the cross-section of commodity returns through a long-short strategy. Each month, commodities are sorted into three groups—High, Mid, and Low—based on media emotion intensity. By going long on the High and short on the Low portfolios, this analysis constructs a High-minus-Low portfolio, which yields an impressive 13.51% (t-stat = 2.95) average annualized return, outperforming the basis-momentum (12.50%) and momentum (9.57%) long-short portfolios. Even when benchmarking against the key factors suggested by the literature, the media emotion intensity factor still stands out, delivering significant abnormal returns of around 11.50% in the spanning test. This analysis further explores the relationship between media emotion intensity and other factors, such as media coverage, basis, momentum, and basis-momentum, by conducting a double sort on commodities. The findings reveal a notable pattern: media emotion intensity significantly enhances returns for commodities with lower media coverage, while the effect is negligible for high media coverage commodities. This emphasizes the stronger influence of media emotion intensity on less media-featured commodities. Moreover, the premium from media emotion intensity is particularly significant for commodities with

backwardation, high momentum, and high basis-momentum.

Second, given the predictive ability of media emotion intensity for commodity futures returns, as evidenced by the significant performance of the long-short strategy, this study proceeds to examine whether media emotion intensity can predict trading behaviors amongst different trader categories. The focus is on two distinct categories: commercial traders, who are typically regarded as hedgers, and non-commercial traders, who are typically regarded as speculators (Kang et al., 2020). Following the methodology adopted by Kang et al. (2020), this research uses the Fama-Macbeth regression framework to study the impact of media emotion intensity on the trading propensities of both commercial and non-commercial traders. The findings indicate that media emotion intensity in the previous month positively correlates with commercial traders' short positions and non-commercial traders' long positions in the current month. This aligns with Kang et al. (2020) that in the short term, commercial traders typically provide liquidity to meet the increased demand from non-commercial traders. Media emotion intensity remains significant even after controlling for news sentiment, past return, and preceding month trading tendencies, indicating its independent influence on trader behavior in the following month.

Third, this study examines the risk premium of media emotion intensity in the cross-section of commodity futures returns. The predictive power of media emotion intensity is examined using the Fama-Macbeth cross-section test against benchmark factors, including basis, momentum, and basis-momentum.

The results of this test, at both the portfolio and the individual commodity levels, indicate a significant and stable premium for media emotion intensity, around 1.20% monthly or 14.40% annually, echoing results from univariate sorting. This suggests that the estimated premium of the media emotion intensity factor is robust and not subsumed by other commonly considered benchmark factors. Further, following Feng et al. (2020), this study conducts LASSO regressions to identify the most relevant factors with the greatest predictive potential. Three LASSO techniques, regular, adaptive, and plug-in, are applied to select the most relevant factors in 200 randomized subsamples. The results show that media emotion intensity is the most selected among the test factors, confirming that media emotion intensity is an important predictor for the cross-section of commodity futures returns.

Finally, this study investigates whether the media emotion intensity premium is subsumed by other risk sources. These risk sources include hedging pressure (Boons and Prado, 2019; De Roon et al., 2000; Kang et al., 2020; Keynes, 1930), inventory (Boons and Prado, 2019; Gorton et al., 2013; Kaldor, 1939), market liquidity and funding liquidity (Boons and Prado, 2019; Brunnermeier and Pedersen, 2009), market volatility (Boons and Prado, 2019; Brunnermeier and Pedersen, 2009), stock market and downside stock market (Boons and Prado, 2019; Brunnermeier and Pedersen, 2009), and macroeconomic risk (Le Pen and Sévi, 2018; Szymanowska et al., 2014). The results of the cross-section tests confirm that the media emotion intensity premium remains significant at around 14% annualized after

controlling for the above risk sources. The robustness of the media emotion intensity premiums highlights that the predictive power of the media emotion intensity is not subsumed by the commonly considered risk sources in the commodity futures market.¹

This paper contributes to the literature on commodity asset pricing and, specifically, the role of media news in the commodity futures market. The research suggests that media emotion intensity is a potentially important new factor that significantly explains the cross-section of returns in the commodity futures market. The average abnormal return of the media emotion intensity factor is comparable to that of the basis-momentum factor in all the main tests over our sample period. Moreover, it is far more significant, both statistically and economically, than all the other benchmark factors. Not only does the media emotion intensity factor survive the time-series spanning test and the Fama-Macbeth cross-section tests, but it also ranks top in the three LASSO regression tests. Additionally, we find that media emotion intensity only generates strong impacts on commodities with low media coverage, backwardation, high momentum, and high basis-momentum. Finally, we show meaningful evidence that media emotion intensity predicts the trading tendencies of commercial and non-commercial traders.

¹Our data sample, commencing from 1998, is constrained by the media news data availability from Thomson Reuters. This shorter time period results in lower t-statistics for factor premiums in our cross-section tests compared to prior studies, such as Boons and Prado (2019), covering longer periods. Consequently, to test the premium of our factor, we only examine the statistical significance of media emotion intensity premium and its comparative t-statistic with those of other benchmark factors.

The rest of our paper is organized as follows. Section II describes the data and methodology. Section III examines the factor premium of media emotion intensity. Section IV investigates the cross-sectional impact of media emotion intensity on trader behaviors. Section V conducts cross-sectional tests and LASSO regressions to further test the risk premium of media emotion intensity. Section VI concludes. A further description of the media data source and discussions is presented in the Internet Appendix.

2. Data and Methodology

2.1. Commodity futures data

This study uses commodity futures data from Barchart[®], formerly known as the Commodity Research Bureau (CRB), which is the same data source utilized in the research conducted by Szymanowska et al. (2014) and Boons and Prado (2019). The data includes 26 commodity futures traded on CBOT, NYMEX, COMEX, CME, and ICE exchanges, spanning from January 1998 to February 2020. The selection of these commodity futures is based on the availability of media emotion intensity data from Thomson Reuters. The sample period starts in January 1998, which aligns with the start of Thomson Reuters' collection of the media data used in this study. Following Boons and Prado (2019), the monthly excess returns on a fully collateralized futures

position are calculated as

$$R_{i,t+1}^{T_n} = \frac{F_{i,t+1}^{T_n}}{F_{i,t}^{T_n}} - 1 \quad (1)$$

where $R_{i,t+1}^{T_n}$ is the return of n^{th} nearby futures contract of the commodity i for the month $t + 1$. $F_{i,t}^{T_n}$ is the price of the n^{th} nearby futures contract of commodity i at the end of month t . This study focuses on the first and second nearby futures contracts due to their higher liquidity compared to contracts that are further from maturity. The approach followed, as described by Boons and Prado (2019), defines the first nearby contract as the contract that can be traded at the end of month t , with an expiration after month $t + 2$, and the closest maturity, denoted as T_1 . This selection avoids the need for rolling considerations and ensures the continuous tradability of the selected futures prices when holding for a month. This time gap is sufficient to avoid any anomalous trading and price movements as the contract approaches its maturity (Boons and Prado, 2019). The second nearby contract is defined as the contract that is immediately following the first nearby contract in terms of maturity, denoted as T_2 , and can be traded at the end of month t .

According to Boons and Prado (2019) and Szymanowska et al. (2014), commodity futures markets have spot premiums and term premiums. Spot premiums are better captured by longing the first nearby futures contract, while term premiums are captured by longing the first nearby contract and shorting the second nearby contract. In this study, the nearby return, R_{t+1}^{nb} ,

represents the return of the first nearby futures contract, and the spreading return, $R_{t+1}^{sp} = R_{t+1}^{T_1} - R_{t+1}^{T_2}$, is the return difference between the first and second nearby contracts. The literature has suggested key factors that can predict the cross-section of commodity futures return. Szymanowska et al. (2014) demonstrate that the basis factor can predict both the nearby and spreading returns. Bakshi et al. (2019) indicate that the momentum factor is priced in the average nearby return. Boons and Prado (2019) introduce the basis-momentum as an additional important factor and demonstrate that the basis, momentum, and basis-momentum factors are the three key tradable factors in the commodity futures market. This study follows Boons and Prado (2019) to construct these factors as follows. First, basis is measured as

$$Basis_{i,t} = \frac{F_{i,t}^{T_2}}{F_{i,t}^{T_1}} - 1 \quad (2)$$

Second, momentum is calculated based on the returns of the past 12 months as:

$$Mom_{i,t} = \prod_{k=t-11}^t (1 + R_{i,k}^{T_1}) - 1 \quad (3)$$

Third, basis-momentum is measured as the difference in the momentum of first and second nearby futures contracts as:

$$BM_{i,t} = \prod_{k=t-11}^t (1 + R_{i,k}^{T_1}) - \prod_{k=t-11}^t (1 + R_{i,k}^{T_2}) \quad (4)$$

2.2. Media news data

This study focuses on a new media factor, which we refer to as media emotion intensity, along with two other well-known media factors: news sentiment and media coverage. These factors are measured using the Thomson Reuters Marketpsych Indices (TRMI) database, which analyzes news content from various financial media news outlets supplied by the Refinitiv Machine Readable News (MRN). TRMI has several advantages, including coverage of a wide range of financial news outlets, separation of content for each asset in each news item, and an anti-correlation process to eliminate word misidentifications, such as "gold" in "gold medal" rather than as a commodity. The TRMI data provides real-time reports with a maximum lag of 500 milliseconds for news items to be fed into the system.

In this study, media emotion intensity quantifies the ratio of emotional content to factual content. It differs from the well-established news sentiment (García, 2013; Smales, 2014; Tetlock, 2007; Tetlock et al., 2008), which relies on the classification of words and phrases in media reports as positive or negative. This classification can be challenged due to the varied sentimental meanings that a single word can convey in different contexts. In contrast, measuring media emotion intensity is more straightforward and transparent as it focuses on differentiating emotional content from factual content. It provides insight into the emotional intensity of media news towards a specific asset. Journalists contribute to news articles by adding words and phrases

that carry emotional content to a collection of factual materials. Therefore, considering the amount of factual content is crucial when calculating the intensity of emotions. Also, if a news item is more emotional in a specific aspect of emotion, it does not necessarily mean it contains more emotional content compared to factual content. Therefore, media emotion intensity in this study only measures the extent to which media adds emotional content to the factual content in news, not the level of any specific emotion.

This study commences with a well-known media factor, media coverage. Media coverage reflects how much content the media news provides readers about a specific asset in a specific period. Fang and Peress (2009) quantify this factor by counting the number of news items referencing each stock. This study uses a variable called "*buzz*" in TRMI data, which may more accurately convey the concept of media coverage. For a certain one-minute period, the system records all words and phrases related to a specific asset a . This set of words is denoted by the symbol $V(a)$. The system then scores each word v in the set $V(a)$. Let $S(v)$ denote the score value of v . The "*buzz*" generated for the asset a over a period of one minute is calculated as

$$buzz_t^{1-min}(a) = \sum_{v \in V(a)} S(v) \quad (5)$$

Given that the system reports news in the Coordinated Universal Time (UTC) zone, all timestamps in the TRMI data are converted to the Central Standard Time (CST) zone. For the month T , the media coverage of the

asset a is calculated as the sum of all one-minute "buzz" for the asset a encompassing the period from the daily settlement time of the last trading day of the month $T - 1$ to the daily settlement time of the last trading day of the month T . The more news content related to the asset a , the more variables are recorded in the system. As a result, "buzz" will take on a higher value.

To measure media emotion intensity, The TRMI data system determines whether each word v in the set of words $V(a)$ related to the asset a refers to emotional or factual content. Specifically, it defines the function $I_{emo}(v)$ as

$$I_{emo}(t, v) = \begin{cases} +1 & \text{if } v \text{ refers to emotional content} \\ -1 & \text{if } v \text{ refers to factual content.} \end{cases} \quad (6)$$

Following this, the media emotion intensity of asset a is constructed as

$$Media \text{ emotion intensity}(a) = \frac{\sum_{v \in V(a)} (I_{emo}(v) \times S(v))}{buzz(a)}. \quad (7)$$

Because this indicator is scaled by "buzz", it takes a value from -1 to 1 and reflects the proportion of purely emotional aspects in the news content. A higher value of the indicator corresponds to a higher intensity of emotional content compared to factual content. In the TRMI data, the one-minute media emotion intensity is provided as the one-minute "EmotionVsFact" indicator. This study calculates media emotion intensity for monthly frequency as the average of one-minute media emotion intensity weighted by one-minute

"buzz". News sentiment is also measured using the same approach as for media emotion intensity. To assist in sorting commodities based on news sentiment, this study uses a net sentiment measure ranging from -1 to 1. When news sentiment is closer to 1, the media is more positive toward the asset. Conversely, when news sentiment is closer to -1, the media is more negative toward the asset². The descriptive statistics of the key factors in this analysis are reported in Table 1.³

[Table 1 here]

Media emotion intensity in this study conveys a distinct meaning compared to news sentiment. While news sentiment reflects the positive or negative tone of media news towards an asset, media emotion intensity indicates the intensity of emotion associated with factual content in the news. Table 2 presents the correlations between media emotion intensity, news sentiment, and media coverage. The table includes five types of correlations: (1) full sample correlation, (2) correlation for observations with positive news sentiment, (3) correlation for observations with negative news sentiment, (4) time-series correlation, and (5) cross-section correlation. To measure time-series correlation, this study calculates the correlation for each commodity in the full sample period and then takes the average of these correlations.

²Internet Appendix A.1 and A.2 provide further discussion on the specifics of the linguistic analysis for TRMI data and the news sentiment measurement, respectively.

³This study also reports the descriptive statistics of the nearby and spreading returns for each commodity futures in Table A1 and the descriptive statistics of media factors for each commodity futures in Table A2 in the Internet Appendix.

To calculate cross-section correlation, this study computes the correlation for each month and then averages the monthly correlations. Table 2 reveals a weak correlation between media emotion intensity and news sentiment, with a maximum magnitude of 0.177 for the correlation when news sentiment is positive. This finding supports the notion that media emotion intensity captures distinct meanings from news sentiment. The correlations between media emotion intensity and media coverage are also relatively low, with a maximum value of 0.303 for time-series correlation. These results suggest that media emotion intensity can provide additional insights beyond the key media factors discussed in the literature.

[Table 2 here]

3. Media emotion intensity and Long-short strategy

3.1. Long-short strategy approach

This section explores the predictability of media emotion intensity for the cross-section of commodity futures returns. The portfolio performance is tested by sorting commodities based on media emotion intensity and forming long-short portfolios for a one-month period. A comparison is made between the performance of the media emotion intensity portfolio and those of other

media and benchmark factors. Specifically, at the end of each month, commodity futures are ranked according to their signal values, and portfolios are constructed accordingly. The High4 portfolio holds the four commodities with the highest signal values, while the Low4 portfolio consists of the four commodities with the lowest signal values. The Mid portfolio includes the remaining commodities with available futures contracts.

Additionally, a High4 minus Low4 portfolio (High4-Low4) is created by taking long positions on the High4 portfolio and short positions on the Low4 portfolio. The performance of these portfolios is evaluated using both nearby and spreading returns. Nearby returns are assessed by taking long positions on the High4-Low4 portfolio using the first nearby commodity futures contracts. Spreading returns involves taking long positions on the High4-Low4 portfolio using the first nearby contracts and short positions using the second nearby contracts. Each portfolio is formed at the end of month t and held until the end of month $t + 1$. Boons and Prado (2019) claim that by examining nearby and spreading returns, the sorting aids in assessing the predictability of factors in the cross-section setting and across maturities.

3.2. Single-sorting results

Table 3 reports the performance of long-short portfolios sorted by media and benchmark factors. Among three media factors and three benchmarks, media emotion intensity has the most significant effects, with an average nearby return of 13.51% and a t-stat of 2.95 for the High4 - Low4 portfolio.

Basis momentum still shows the highest effect among the benchmarks, with an average nearby return of 12.50% and a t-stat of 2.66. However, the effect of basis-momentum in our sample is weaker compared to the sample from August 1960 to February 2014 in Boons and Prado (2019). These results suggest that media emotion intensity can be a potential predictor for commodity futures returns, with a higher average nearby return and t-stat compared to basis-momentum. Media coverage may also be a good candidate, with a high average nearby return for the Low4 - High4 portfolio (9.16% with a t-stat of 2.69). The effect of media coverage is consistent with that reported in Fang and Peress (2009) for the equity market.

[Table 3 here]

For spreading returns, basis momentum still shows the greatest effect, with an average spreading return of 4.64% (t-stat = 4.78) for the High4 - Low4 portfolio. Interestingly, both media emotion intensity and media coverage can significantly predict spreading returns, with average spreading returns of 3.35% (t-stat = 4.97) and -2.11% (t-stat = -3.88) for the High4 - Low4 portfolio, respectively. This analysis also examines the predictability of news sentiment in the cross-section setting. However, the findings suggest that news sentiment does not significantly predict the cross-section of commodity futures returns using monthly frequency. Therefore, the rest of this study will focus solely on media emotion intensity and media coverage.

According to Boons and Prado (2019), the transaction costs associated

with the basis-momentum long-short strategy could be as high as 158 basis points. When sorting by media emotion intensity, it is observed that precious metals, crude oil, natural gas, and orange juice are regularly featured in the High4 portfolios, while agricultural commodity futures are frequently included in the Low4 portfolios ⁴. The estimated transaction costs of 158 basis points imply that the 13% returns of the High4-Low4 portfolio ranked by media emotion intensity can comfortably exceed the cost. Similarly, the 9.16% profit of the High4-Low4 portfolio sorted by media coverage also survives the transaction cost.

One concern about portfolio formation based on media emotion intensity is that the average of this factor may be influenced by the nature of the commodity. However, statistics on the frequency of each commodity in the High4 and Low4 groups show that no commodity is consistently present in the group with the highest or lowest media emotion intensity ⁵. Precious metals have the highest proportion of occurrences in the High4 group. This is similar to the equity market, wherein some prominent firms recurrently appear in the upper quantile when sorting by firm size. However, most commodities have occasionally appeared in both the High4 and Low4 groups.

To address a potential bias in the results due to certain commodities

⁴Table A3 in the Internet Appendix shows the percentage of times that each commodity futures appears in the High4 and Low4 portfolios.

⁵Table A3 in the Internet Appendix provides descriptive information on the percentage of each commodity appearing in the High4 and Low4 groups, sorted by media and benchmark factors.

appearing frequently in the High4 and Low4 portfolios, this study examines the performance of the High6-Low6 and High8-Low8 portfolios. The High k -Low k portfolios involve taking long positions in the k commodity futures with the highest factor values and short positions in the k commodity futures with the lowest factor values. The performance of these portfolios is reported in Table 4. The results indicate that both the High6-Low6 and High8-Low8 portfolios, sorted based on media emotion intensity and media coverage, yield significant average nearby returns at around 10.2% and 7.19%, respectively. To further enhance the robustness of the analysis, commodities that appear more than 70% in the High4 and Low4 portfolios are eliminated. The High4-Low4 portfolio is then reconstructed using narrower sets of commodities. Table 4 shows that after removing commodities that appear most frequently in the groups of commodities with the highest and lowest media emotion intensity values, the High-Low4 portfolio still generates significant average nearby returns.

[Table 4 here]

This analysis also examines the overlap between portfolios sorted by media factors and benchmark factors. It calculates the percentage of common commodities between each pair of portfolios among High4 and Low4 portfolios sorted by media and benchmark factors⁶. The portfolios sorted by media emotion intensity have a weak overlap with those sorted by other factors, with

⁶The results are presented in Table A4 of the Internet Appendix.

a maximum overlap of 34% observed with the High4 portfolio sorted by news sentiment. This analysis suggests that the returns of the long-short portfolio, sorted by media emotion intensity, are not strongly influenced by news sentiment, media coverage, basis, momentum, or basis-momentum.

3.3. Time-series spanning test

This study employs the time series spanning test to determine whether the return of portfolios sorted by other benchmark factors span the return of the portfolio sorted by media emotion intensity. Boons and Prado (2019) conduct the spanning test for basis-momentum against the factor set proposed by Szymanowska et al. (2014) and Bakshi et al. (2019). The Boons and Prado (2019) spanning test highlights the importance of testing the significance of alpha in assessing the additional value of a factor, which is discussed in Barillas and Shanken (2017, 2018). We follow the same logic as Boons and Prado (2019).

The results of the spanning tests are summarized in Table 5. This table examines whether the returns to the portfolio sorted by the media emotion intensity are spanned by two sets of factors proposed in (1) Bakshi et al. (2019) and Boons and Prado (2019), and (2) Szymanowska et al. (2014) and, again, in Boons and Prado (2019). Szymanowska et al. (2014) propose three benchmarks: the nearby return of the High4 - Low4 portfolio sorted by basis (basis [nb]), as well as the spreading returns of the High4 and Low4 portfolios also sorted by basis (basis [high4-sp] and basis [low4-sp], respectively).

Subsequently, Bakshi et al. (2019) introduce two more benchmarks, namely, the average return of all commodity futures weighted equally (average [nb]) and the momentum factor (momentum [nb]) calculated based on the nearby return of the High4 - Low4 portfolio sorted by momentum. Later, Boons and Prado (2019) introduce the other two factors, which are the nearby and spreading returns of the High4 - Low4 portfolio sorted by basis-momentum (basis-momentum [nb] and basis-momentum [sp], respectively). To determine whether the addition of media factors improves mean-variance efficiency, we apply Gibbons et al. (1989) joint GRS test for both the nearby and spreading returns of each factor.

[Table 5 here]

Table 5 shows that even after considering various factors such as basis-momentum, basis, momentum, average factor, and media coverage, the alpha of the nearby portfolio sorted by high media emotion intensity remains significant at 11.35% (t-stat = 2.30) in Model 1. Model 3, with a different set of factors, also confirms these findings, with an alpha of 11.72% (t-stat = 2.26). The R-squared values for media emotion intensity are low in both Models 1 and 3 (0.09 and 0.04, respectively). This suggests that the performance of the media emotion intensity nearby portfolio is only slightly correlated with other benchmark factors, emphasizing the usefulness of the media emotion intensity factor. The table also reports significant alphas for the spreading portfolio of media emotion intensity in Models 2 and 4. However, the alphas

for the spreading portfolio are small, around 2%. Following these spanning tests, the GRS tests for the media emotion intensity portfolio yield large F-stats with p-values below 1%. These results demonstrate that media emotion intensity can greatly improve mean-variance efficiency.

3.4. Double-sorting results

In this section, the investigation focuses on whether media emotion intensity generates additional returns for portfolios sorted by other factors. To accomplish this, commodities are double-sorted based on media emotion intensity and each factor independently. Independent sorting is conducted through the separate ranking of commodities according to media emotion intensity and each of the factors. Specifically, the commodities are sorted into High and Low portfolios based on the median value of media emotion intensity for each month. Additionally, the commodities are divided into portfolios based on other factors: (1) contango and backwardation, determined by the positive or negative value of the basis in month t , and (2) high and low, based on the median value of the factor in month t . News sentiment is excluded from the analysis as the High4-Low4 portfolio, ranked by news sentiment, does not yield a significant average return.

Table 6 presents the results of independent sorting on media emotion intensity and other factors. The third and fourth columns display average returns and t-stat for testing single-sorted portfolios by each factor. The High-Low portfolio consists of long positions in commodities with high factor

value and short positions in commodities with low factor value. When single-sorting commodities into two levels, the High-Low portfolio sorted by media emotion intensity generates a 6.04% return on average with a t-stat of 2.10. Compared to High4-Low4 portfolios, the profit of the High-Low portfolio sorted by media emotion intensity is significantly lower, indicating that the commodities with the highest and lowest values of media emotion intensity may have the most extreme returns.

[Table 6 here]

Examining media coverage, the Low portfolio generates a higher average return than the High portfolio. This result is consistent with Fang and Peress (2009) that assets with higher media coverage earn higher returns. However, in the commodity futures market and during our sample period, both the High and Low portfolios sorted by media attention generate statistically insignificant average returns. Within the group of high media coverage commodities, it is interesting that both high and low media emotion intensity portfolios earn similar average returns at around 0.12% and are not statistically significant. The High-Low portfolio sorted by media emotion intensity within high media coverage commodities generates only 0.01% (t-stat = 0.00) return on average. This result suggests that when assets are highly featured in the media, media emotion intensity has negligible effects on investors and asset returns. However, among commodities with low media coverage, the High portfolio ranked by media emotion intensity produces

a significant 9.17% average return ($t\text{-stat} = 2.34$), while the Low portfolio generates a negative 3.53% average return ($t\text{-stat} = -1.11$). The High-Low portfolio sorted by media emotion intensity within the group of low media coverage commodities earns a return of 12.70% ($t\text{-stat} = 3.39$) on average, a huge improvement from the 2.42% return of the Low portfolio sorted by media coverage. This suggests that media emotion intensity has a significant impact on the return prediction for commodities with low media coverage.

In single sorts, basis, momentum, and basis-momentum generate significant returns of 8.11%, 6.88%, and 9.27%, respectively. When separating the high and low media emotion intensity commodities within each portfolio sorted by these factors, similar results are obtained, with the High portfolios sorted by media emotion intensity producing higher average returns than the Low portfolio. The High-Low portfolio sorted by media emotion intensity improves the average return of backwardation commodities from 7.87% ($t\text{-stat} = 1.86$) to 13.90% ($t\text{-stat} = 2.54$), the average return of high momentum commodities from 4.94% ($t\text{-stat} = 1.48$) to 7.67% ($t\text{-stat} = 2.07$), and the average return of high basis-momentum commodities from 6.00% ($t\text{-stat} = 1.77$) to 8.85% ($t\text{-stat} = 2.38$). The returns of the above High-Low portfolios sorted by media emotion intensity are mainly derived from the high media emotion intensity portfolios, with averages of 11.53% ($t\text{-stat} = 2.34$), 8.61% ($t\text{-stat} = 2.22$), and 9.98% ($t\text{-stat} = 2.54$) for basis, momentum, and basis-momentum respectively. Overall, Table 6 reveals that the impact of media emotion intensity is stronger for commodities with low media coverage, backwardation,

high momentum, and high basis-momentum and weaker for those with high media coverage, contango, low momentum, and low basis-momentum.

4. Media emotion intensity and Trading behaviors

The performance of the long-short strategy, which is sorted based on the intensity of media emotion, has indicated that media emotion intensity can predict the cross-section of commodity futures returns. The spanning test result also suggests that the return of the High4-Low4 portfolio, which is sorted based on media emotion intensity, is only weakly explained by media coverage and other benchmark factors. These findings imply that media emotion intensity may possess the ability to predict the trading behaviors of traders in the commodity futures market. Therefore, this section aims to investigate whether the intensity of media emotions can predict the tendency of traders to engage in long or short trading.

To facilitate this examination, the research uses trader position data from the Commodity Futures Trading Commission (CFTC). The CFTC regularly publishes weekly reports known as Commitment of Trader reports, which provide a summary of long and short positions held by commercial and non-commercial traders in the commodity futures market. These reports are typically released on Fridays and cover data for the preceding Tuesday-to-Tuesday period. Aggregating weekly data into monthly data poses chal-

lenges, and in this study, a week (from Tuesday to Tuesday) is assigned to a specific month if the final day of that week falls within that month. In order to capture the dynamics of buying and selling tendency, this analysis quantifies the disparity between long and short positions for each trader category and normalizes this difference by open interest. Specifically, this study defines the shorting pressure exerted by commercial traders and the longing pressure exhibited by non-commercial traders as follows:

$$\text{Commerical shorting pressure}_t = \frac{S_{c,t} - L_{c,t}}{OI_t} \quad (8)$$

$$\text{Non-commerical longing pressure}_t = \frac{L_{nc,t} - S_{nc,t}}{OI_t} \quad (9)$$

As commercial/non-commercial traders are commonly known as hedgers/speculators in the commodity futures market, this study emphasizes the shorting/longing tendency of commercial/non-commercial traders.⁷

4.1. Predictability of media emotion intensity

This analysis uses the Fama-Macbeth cross-section regression method, following a similar approach as in Kang et al. (2020), to examine the cross-section impact of media emotion intensity on shorting/longing pressure of commercial/non-commercial traders. Each month, a cross-section regression

⁷The shorting pressure exhibited by commercial traders follows a similar formula as hedging pressure. However, because of this study's focus on the short-term impact of media emotion intensity on trading tendency, this metric is referred to as shorting pressure to distinguish it from hedging pressure, which reflects the long-term shorting behavior of hedgers.

is conducted to analyze the impact of media emotion intensity and other control factors on the shorting/longing pressure of commercial/non-commercial traders. Subsequently, the time-series average of the cross-sectional effects is calculated, and the statistical significance of these average effects is tested. The findings are presented in Table 7, with the numbers in parentheses representing the Newey West corrected t-statistic for the average effect.

[Table 7 here]

We employ five models to evaluate the effects of commercial shorting and non-commercial longing pressure. Model 1 incorporates only the first lag of media emotion intensity. Model 2 integrates the first lags of news sentiment and return. Model 3 includes all variables from Models 1 and 2 to assess the extra predictive power of media emotion intensity beyond news sentiment and return's first lags. Model 4 supplements Model 2 with the first lags of basis, momentum, and basis-momentum. Finally, Model 5 examines the additional predictive power of media emotion intensity over the factors in Model 2.

Model 1 reveal that commodities with higher media emotion intensity in the previous month are significantly associated with increased shorting pressure from commercial traders. Even after considering news sentiment and past return in the model, the average cross-section impact of media emotion intensity remains statistically significant at around 0.540 (t-stat = 12.110). Notably, the coefficient of media emotion intensity only decreases

slightly from Model 1 to Model 3, indicating that the predictability of emotion intensity on the shorting pressure of commercial traders is not strongly correlated with the impact of news sentiment and past commodity futures return. Incorporating media emotion intensity into Model 2 increases the adjusted R-squared from 0.0119 to 0.194. When adding further basis, momentum, and basis-momentum as predictors in Model 2 to form Model 4, adding media emotion intensity to Model 4 improves the adjusted R-squared from 0.156 to 0.223. The coefficients of media emotion intensity in Model 5 (0.517) are similar to those in Model 3 (0.540), indicating a weak connection between the predictability of media emotion intensity and that of basis, momentum, and basis-momentum.

Table 7 also demonstrates the significant predictive ability of media emotion intensity for the cross-section of non-commercial traders' longing pressure. Specifically, in the presence of high media emotion intensity, non-commercial traders exhibit a stronger inclination towards longing compared to shorting. The inclusion of news sentiment and past return in the model slightly reduces the magnitude of the impact of media emotion intensity, from 0.360 (t-stat = 10.863) to 0.277 (t-stat = 7.880). This finding indicates a weak relationship between the cross-section predictability of media emotion intensity and those impacts of news sentiment and past return. When we expand the factor set to include additional factors like basis, momentum, and basis-momentum, the media emotion intensity still enhances the adjusted R-squared value. It increases from 0.159 in Model 4 to 0.192 in Model 5. The

coefficient of media emotion intensity in Model 5 (0.228) is similar to that in Model 3 (0.277). This suggests that the predictability of the media emotion intensity for non-commercial longing pressure has a weak association with factors like basis, momentum, and basis-momentum.

4.2. Media coverage's moderating effect

The results of the double sorting analysis in Table 6 reveal that the average return of the High-Low portfolio, which is sorted based on media emotion intensity, is only significant for commodities with low media coverage, but not significant for those with high media coverage. This suggests that media emotion intensity may have a greater impact on commodities that are less covered by the media. To examine the moderating effect of media coverage on the relationship between media emotion intensity and trading behavior, this study introduces an interactive term between media emotion intensity and a dummy variable representing commodities with low media coverage into the model. Specifically, each month, the commodities are divided into two groups based on the median value of media coverage: high media coverage and low media coverage. Subsequently, the Fama-Macbeth cross-section test is used to regress the following model:

$$\begin{aligned}
Y_{i,t+1} = & \alpha_t + \beta_t EMOTION_INTENSITY_{i,t} \\
& + \gamma_t EMOTION_INTENSITY_{i,t} \times LOW_MEDIA_COVERAGE_{i,t} \quad (10) \\
& + \theta_t NEWS_SENTIMENT_{i,t} + \phi_t RETURN_{i,t} + \rho_t Y_{i,t} + u_{t+1}
\end{aligned}$$

The shorting pressure of commercial traders and the longing pressure of non-commercial traders at the end of each month are denoted as $Y_{i,t}$, representing the respective trading behaviors for commodity i . The results of Table 8 exhibit the moderating effect of media coverage. In all models, the estimated coefficients of the interactive term are positive and statistically significant. The interactive term between emotion intensity and media coverage remains significant at around 0.315 (t-stat = 15.071) and 0.283 (t-stat = 11.452) when controlling for news sentiment and past return in Models 3 and further basis, momentum and basis-momentum in Model 7, respectively. These findings suggest that the cross-sectional impact of media emotion intensity is more pronounced for commodities with low media coverage.

[Table 8 here]

Furthermore, this moderating effect remains significant at around 0.283 (t-stat = 11.452) and 0.166 (t-stat = 8.711) even when considering the lag of basis, momentum, and basis-momentum in Models 3 and 6, respectively. Consequently, it implies that the heightened predictability of media emotion intensity on subsequent month trading tendencies for low media coverage

commodities is not encompassed within the impact of basis, momentum, and basis-momentum. The findings from Table 8 provide evidence to suggest that the impact of media emotion intensity on traders is more pronounced when there is limited information available about the asset. This supports the view that media emotion intensity has a stronger impact on traders' behavior in situations where they have less knowledge about the asset.

5. Media emotion intensity and the cross-section of commodity futures returns

Given the significant performance of the long-short portfolio sorted by media emotion intensity and the cross-section impact of media emotion intensity on the trading tendency of traders in the commodity futures market, the next purpose of this study is to test whether the investors are compensated for investing in higher emotion intensity commodities or whether media emotion intensity explains the cross-section of commodity futures returns. To achieve this, this study constructs two factors: "Emotion intensity (nb)" and "Emotion intensity (sp)", which are the nearby return and the spreading return of the High4 - Low4 portfolio sorted by media emotion intensity, respectively. Then, these two factors are used in the two-stage Fama-Macbeth cross-section regression for commodity futures returns with other benchmark factors to test the premium of media emotion intensity.

The relevant literature has suggested several benchmark factors for com-

modity futures returns, as outlined in Table 5. Boons and Prado (2019) argue that little is known about the cross-section of commodity futures. Thus far, the most popular factors are those given by Szymanowska et al. (2014), Bakshi et al. (2019), and, most notably, Boons and Prado (2019). These factors, which are previously used in the spanning tests, are also incorporated in this cross-sectional analysis of media emotion intensity. This study employs the two-stage Fama-Macbeth regression to conduct the cross-section test at two levels: portfolio and individual commodity levels.

5.1. Cross-section test at the portfolio level

This analysis begins with a cross-section test at the portfolio level. Firstly, the commodities are divided into three portfolios based on their basis, momentum, basis-momentum, media emotion intensity, and media coverage, resulting in 15 portfolios. After segregating nearby and spreading returns, a total of 30 portfolio return series are obtained. Secondly, the commodities are categorized into six groups: energy, grains, metals, soft commodities, food oils, and livestock. This classification generates 12 portfolio return series by considering both the nearby and spreading returns of these six commodity-type portfolios. In total, there are 42 portfolio test return series. The two-stage cross-section tests at the portfolio level begin by regressing each portfolio return series against the factor set in each model to determine the exposures of the portfolios to the factors. This time-series regression covers the entire sample period from February 1998 to February 2020. Hence,

the exposures (betas) of portfolios in the first stage are time-invariant. In the second stage, the returns of the portfolios are regressed on their exposures for each month to estimate the monthly premium. The risk premium of each factor is then calculated as the average of the monthly premiums. The objective of this analysis is to examine whether portfolios that exhibit greater sensitivity to the media emotion intensity factor over a long time are rewarded with higher average returns.

Table 9 presents the results of the portfolio-level cross-section tests. The estimated premium for each factor is shown, along with the Newey-West corrected t-statistic in parentheses. The standard errors are further adjusted for the errors-in-variables problem in the first-stage regression, following the method proposed by Shanken (1992). The adjusted t-statistics according to Shanken (1992) are displayed in square brackets below the Newey-West t-statistic. The first three models analyze the factors from Bakshi et al. (2019) and Boons and Prado (2019), while the last three models examine the results for the factors from Szymanowska et al. (2014) and Boons and Prado (2019). The expectation is that the t-statistics in our test will not be as high as those in Boons and Prado (2019) due to our shorter sample period.

[Table 9 here]

The emotion nearby factor consistently exhibits a premium of approximately 1.2% per month (equivalent to around 14.5% per year) across the test models. All estimated premiums for this factor are statistically signif-

icant, with Shanken t-statistics greater than 2. These findings suggest that the premium associated with media emotion intensity is significant in the cross-section of tradable portfolio returns. Furthermore, the premium of this factor is slightly higher than that of basis-momentum, averaging around 1.1% per month (13% per year). The consistency of the estimated premium and the t-statistics for the basis-momentum nearby factor indicate that its significance is unlikely to be connected to media emotion intensity. No significant premium is observed for the media emotion intensity spreading factor, as the corresponding Shanken t-statistics hover around 1. The intercepts in the last two models with media emotion intensity factors are statistically significant, but they only amount to around 0.1%, contributing approximately 1.2% per year to portfolio returns.⁸

5.2. Cross-section test at the individual commodity level

This study conducts a two-stage Fama-Macbeth regression to test whether media emotion intensity can predict the cross-section of individual commodity returns. Following the methodology outlined in Boons and Prado (2019), we first regress the nearby and spreading returns of commodity futures on the factor set, using a rolling time series regression with a 250-day look-

⁸Additional cross-section tests are conducted for the media coverage factor, and the results are reported in Table A5 in the Internet appendix. However, no significant evidence is found for the premium of the media coverage nearby factor. This suggests that media coverage has limited ability to explain the cross-section of tradable portfolio returns.

back. This approach allows for the estimation of time-varying exposures of assets to factors. In the second stage, the individual commodity nearby and spreading returns are regressed on their respective exposures to the factors. The purpose of this test is to address the arguments made by Lewellen et al. (2010) and Ang et al. (2020) regarding the necessity of examining factors at the individual asset level rather than at the portfolio level. The one-year window for daily returns is chosen instead of the six-year window for monthly returns, as suggested by Boons and Prado (2019), in order to estimate factor betas in a more timely manner.

Table 10 reports the results of the cross-section test at the commodity level. In this panel, we first show the estimated premium for the factors used in each model. There are two values of t-stat displayed for each estimation. First, we correct standard errors to address autocorrelation using the Newey West method with one lag. We report the t-stat for Newey West standard error in the parenthesis under the estimated risk premium. We also correct standard errors using Shanken (1992) method. The Shanken t-stat is reported in the square brackets.

[Table 10 here]

The results of the cross-section test at the commodity level indicate that media emotion intensity exhibits explanatory power in all test models. Across these models, the estimated premiums for this factor consistently average around 1.1% to 1.3% per month (approximately 13% to 15% per

year). These estimations are statistically significant based on the Shanken t-statistic. However, the predictability of media emotion intensity at the individual commodity level is slightly lower compared to the basis-momentum factor. Additionally, the significance of the media emotion intensity nearby premium is weaker than that of the basis-momentum factor. No significant premiums are observed in Models 3 and 6 in Table 10 for the media emotion intensity spreading factor.⁹ Overall, the results suggest that only basis-momentum and media emotion intensity survive the cross-section test at the commodity level.

5.3. LASSO regression

In this study, the predictive power of basis and momentum decreases when other factors are included. This suggests that a factor's ability to explain returns may vary depending on the sample or factors used. According to Feng et al. (2020), a two-pass regression with a LASSO technique is proposed to select relevant factors for cross-section tests in the equity market. Their method identifies primary factors that contribute to explaining returns. Due to the limited number of benchmark factors, cross-section tests with different specifications can be conducted for the commodity futures market. The same

⁹This study also investigates whether the premium of media coverage is present in the cross-section of returns for individual commodity futures. The findings, as reported in Table A6 in the Internet appendix, indicate inconsistent evidence for the predictability of media coverage across the different models. Media coverage premium is found to be significant in explaining the cross-section of individual commodity returns in only two out of the four models examined.

LASSO regression, as suggested by Feng et al. (2020), is used to select factors with the highest predictive ability for returns. To test factor performance consistency across different periods and portfolios, 200 samples are randomly selected from the dataset. Each sample creates a unique set of preferred factors for the cross-section test. The frequency of factor selection indicates its favorability in the test. This study employs three LASSO techniques: standard LASSO, adaptive LASSO (Zou, 2006), and plug-in LASSO (Belloni et al., 2012). Table 11 displays the frequencies of selected factors in the test for 200 randomized samples and three LASSO models.

[Table 11 here]

The standard LASSO model consistently selects the emotion intensity (nb) factor in 100% of the samples. The adaptive LASSO model also shows a high selection rate of 100% for the emotion intensity (nb) factor. The second most frequently chosen factor is the basis momentum, selected 53.5% of the time. Both types of LASSO regression indicate that emotion is the most selected predictor of cross-section portfolio returns. The LASSO plug-in will determine the factors that best predict out-of-sample returns. The emotion intensity (nb) factor is again frequently selected at 87.5%, along with basis-momentum at 82% of the samples. The plug-in LASSO result shows that emotion intensity (nb) and basis-momentum (nb) are the best predictors for portfolio-level cross-section returns, followed by basis (nb) and momentum (nb). These factors align with the benchmark factors mentioned in Boons

and Prado (2019), with the addition of the media emotion intensity factor. In summary, the LASSO test confirms the effectiveness of media emotion intensity factors in explaining commodity futures returns.

5.4. Media emotion intensity premium and other risk sources

Some hypotheses have been put forward in the literature to explain premium on the commodity futures market. Given the statistical significance of the media emotion intensity premium in the cross-section test with the benchmark factors, we are interested in testing whether other risk sources subsume this premium. To facilitate these tests, we also employ the two-state Fama-Macbeth regression for the intensity of media emotions and each factor. We test the survival of the media emotion intensity premium against hedging pressure, inventory, market volatility, market liquidity, equity market risk, and macroeconomic factors. In this subsection, we focus only on reporting the consistency of the media emotion intensity premium across the tests and leave a detailed discussion of each risk source and corresponding test in Appendix A3.

As the double-sorting results suggest a possible link between hedging pressure and media emotion intensity, we first check to see if hedging pressure can explain the media emotion intensity premium. The predictability of the hedging pressure for commodity futures prices is backed up by the theory of normal backwardation (Keynes (1930)). Accordingly, hedgers trade in the

futures market to hedge for spot price fluctuation. We follow Kang et al. (2020) to use the 12-month moving average hedging pressure in the cross-section test to reflect the true hedging pressure in the market. We begin with the time-series spanning test following the Boons and Prado (2019) approach to regress the return of the High4-Low4 portfolio sorted by media emotion intensity on the High4-Low4 portfolios sorted by hedging pressure and other benchmarks. When adding hedging pressure to the model, the alpha decreases from around 12% annualized to 7.74%, suggesting that the return of the media emotion intensity portfolio is partly related to hedging pressure. However, this abnormal return of 7.74% is still significant and large enough. The result implies that a large proportion of the media emotion intensity portfolio return is not explained by hedging pressure and other benchmark factors. In the cross-section test, the premium of media emotion intensity is statistically significant at 1.2% a month when controlling for the hedging pressure and other benchmarks. This level of premium is similar to the result in the model with only the benchmark factors. This evidence shows that the media emotion intensity premium is not subsumed by hedging pressure.

The storage theory (Kaldor, 1939) suggests a link between commodity inventory level and commodity price. Later, Gorton et al. (2013) argue that the commodity futures premium is driven by the physical inventory level. According to Gorton et al. (2013) and Boons and Prado (2019), the high inventory level is associated with contango, low momentum, and low

basis momentum. Hence, in this study, we use basis, momentum, and basis-momentum to represent the inventory level of commodities. By doing this, we come back with the result of the cross-section test with benchmark factors. As the premium of media emotion intensity is statistically significant, around 1.2% a month in the test, we conclude that the premium of media emotion intensity is less likely to be subsumed by the inventory risk.

According to Boons and Prado (2019), market volatility risk is priced in the cross-section of commodity futures returns. The authors establish that market volatility could predict and explain the basis-momentum component. The negative risk premium associated with market volatility shows that investors are willing to pay a premium to compensate for increased volatility. In Appendix A3.3, we present the cross-section test for market and average volatility. Market volatility and average volatility are not significantly priced in our sample's cross-section of commodity futures returns. However, the Shanken t-stat for market volatility is -1.42, close to significance at the 10% level. Notably, when other benchmark variables are added to the model, the risk price of market volatility is reduced to -0.4%, with a t-statistic of -0.09, meaning that almost all market volatility risk is captured by the risk of media emotion intensity and other benchmark factors. Across the models, the media emotion intensity factor is still statistically significant at around 1.2%.

The literature also suggests that market liquidity and funding liquidity can explain the premium of commodity futures (Boons and Prado, 2019;

Brunnermeier and Pedersen, 2009. We follow Boons and Prado (2019) to use Ted spread as a proxy for funding illiquidity and use VIX as a proxy for market illiquidity. The cross-section test shows that the risks of Ted spread and VIX are significantly priced in the cross-section of commodity futures returns. When adding media emotion intensity into the model, the risk prices of these two liquidity factors lose their significance. However, the premium of media emotion intensity is still statistically significant at around 1.2 a month, suggesting that the market and funding liquidity risk cannot subsume the media emotion intensity premium.

Additionally, Boons and Prado (2019) argue that market risk and downside market risk from the equity market are priced in the cross-section of portfolio returns. If the market returns less than the average minus one standard deviation for the entire sample, the market is considered to be on the downside. A bear equity market may deter speculators from clearing the futures market. We use the excess return on the value-weighted CRSP portfolio to approximate the stock market return. After taking into account the intensity of the emotions in the media in the testing models, the equity market's risk premium is reduced to 2% and is no longer significant (t-stat = 1.20). Similarly, the risk price of the downside equity market also drops to 1.5% (t-stat = 1.4). This conclusion shows that media emotion intensity may partly reflect equity market risk and downside market risk. However, similar to the previous test, the premium associated with the media emotion intensity factor is relatively consistent at 1.2% each month and statistically

significant.

Given that the economic situation affects the supply and demand of commodities, macroeconomic factors are closely related to the commodity market (Le Pen and Sévi, 2018; Szymanowska et al., 2014). Owing to the distinct characteristics of the different commodities, each commodity may react differently to macroeconomic shocks. We collect 10 macroeconomic variables that reflect various facets of the United States economy. Although several of these factors exhibit high correlations, they are not identical. We begin by analyzing these macro variables' principal components to isolate the primary information-containing elements. Factor 1 is mainly connected with CPI, PPI, and M2 supply. Consumer sentiment and unemployment are reflected in Factor 2. Factor 3 is associated with the M1 supply and the US dollar index, which measures the strength of the US dollar. Only the component associated with consumer confidence and the labor market can significantly explain our sample's cross-section of portfolio returns. When the media emotion intensity factor is included in the model, macroeconomic variables' risk prices and the significance of macroeconomic variables are significantly reduced. The premium associated with media emotion intensity again remains statistically significant at around 1.2%. This finding supports the hypothesis that media emotion intensity remains a reliable and consistent predictor of commodity futures returns.

6. Conclusion

This research examines the role of media emotion intensity in the cross-section of commodity futures returns. Media emotion intensity represents a different aspect of media when compared to other common factors, such as news sentiment and media coverage. In our research, the media emotion intensity measures the proportion of emotional content compared to factual content in media news, scaled by the total coverage. This factor, due to the emphasis on the ratio of emotional to factual content in media news, offers a unique perspective on media emotion. It diverges from the examination of specific emotions such as fear or joy.

Our study shows that media emotion intensity has predictive power for commodity futures returns. First, a long-short portfolio sorted by media emotion intensity generates an average annualized return of 13.51% (t-stat = 2.95). Second, the performance of the long-short portfolio sorted by media emotion intensity is weakly correlated with other benchmark portfolios. Third, this study finds that the impact of media emotion intensity on commodity futures return is more profound for commodities with low media coverage, backwardation, high momentum, and high basis-momentum. Fourth, this study demonstrates that trader behavior in the following month is influenced by media emotion intensity, beyond the impact of news sentiment, momentum, and prior trading tendencies.

Moreover, the cross-section Fama-Macbeth tests support the statistical

and economic significance of the premium associated with media emotion intensity, even when controlling for other commonly considered benchmark risk factors. In our robustness check, various LASSO tests rank media emotion intensity as the most significant factor among the benchmark factors considered. Additionally, this research also considers other risk channels such as hedging pressure, funding liquidity, market liquidity, macroeconomic risks, market volatility, average volatility, inventory, equity market risk, and downside equity market risk. However, none of these factors are found to be able to explain the premium of media emotion intensity. The media emotion intensity premium remains statistically significant at around 1.2% per month or approximately 14% per year across cross-section tests with benchmark factors and other risk sources. Media emotion intensity is a consistent and reliable predictor for the cross-section of commodity futures returns.

References

- Adams, Z., & Glück, T. (2015). Financialization in commodity markets: A passing trend or the new normal? *Journal of Banking & Finance*, 60, 93–111.
- Ang, A., Liu, J., & Schwarz, K. (2020). Using Stocks or Portfolios in Tests of Factor Models. *Journal of Financial and Quantitative Analysis*, 55(3), 709–750.
- Bakshi, G., Gao, X., & Rossi, A. G. (2019). Understanding the Sources of Risk Underlying the Cross Section of Commodity Returns. *Management Science*, 65(2), 619–641.
- Barillas, F., & Shanken, J. (2017). Which Alpha? *The Review of Financial Studies*, 30(4), 1316–1338.
- Barillas, F., & Shanken, J. (2018). Comparing Asset Pricing Models. *The Journal of Finance*, 73(2), 715–754.

- Beckett, C. (2015). How journalism is turning emotional and what that might mean for news. Retrieved December 4, 2023, from <https://blogs.lse.ac.uk/polis/2015/09/10/how-journalism-is-turning-emotional-and-what-that-might-mean-for-news/>
- Belloni, A., Chen, D., Chernozhukov, V., & Hansen, C. (2012). Sparse Models and Methods for Optimal Instruments With an Application to Eminent Domain. *Econometrica*, 80(6), 2369–2429.
- Belousova, J., & Dorfleitner, G. (2012). On the diversification benefits of commodities from the perspective of euro investors. *Journal of Banking & Finance*, 36(9), 2455–2472.
- Boons, M., & Prado, M. P. (2019). Basis-Momentum. *The Journal of Finance*, 74(1), 239–279.
- Brunnermeier, M. K., & Pedersen, L. H. (2009). Market Liquidity and Funding Liquidity. *Review of Financial Studies*, 22(6), 2201–2238.
- Da, Z., Engelberg, J., & Gao, P. (2015). The Sum of All FEARS Investor Sentiment and Asset Prices. *Review of Financial Studies*, 28(1), 1–32.
- De Roon, F. A., Nijman, T. E., & Veld, C. (2000). Hedging Pressure Effects in Futures Markets. *The Journal of Finance*, 55(3), 1437–1456.
- Fama, E. F., & MacBeth, J. D. (1973). Risk, Return, and Equilibrium: Empirical Tests. *Journal of Political Economy*, 81(3), 607–636.
- Fang, L., & Peress, J. (2009). Media Coverage and the Cross-section of Stock Returns. *The Journal of Finance*, 64(5), 2023–2052.
- Feng, G., Giglio, S., & Xiu, D. (2020). Taming the Factor Zoo: A Test of New Factors. *The Journal of Finance*, 75(3), 1327–1370.
- García, D. (2013). Sentiment during Recessions. *The Journal of Finance*, 68(3), 1267–1300.
- Gibbons, M. R., Ross, S. A., & Shanken, J. (1989). A Test of the Efficiency of a Given Portfolio. *Econometrica*, 57(5), 1121–1152.
- Goldenberg, A., & Gross, J. J. (2020). Digital Emotion Contagion. *Trends in Cognitive Sciences*, 24(4), 316–328.
- Gorton, G. B., Hayashi, F., & Rouwenhorst, K. G. (2013). The Fundamentals of Commodity Futures Returns. *Review of Finance*, 17(1), 35–105.
- Griffith, J., Najand, M., & Shen, J. (2020). Emotions in the Stock Market. *Journal of Behavioral Finance*, 21(1), 42–56.
- Han, S., Lerner, J. S., & Keltner, D. (2007). Feelings and Consumer Decision Making: The Appraisal-Tendency Framework. *Journal of Consumer Psychology*, 17(3), 158–168.
- Kaldor, N. (1939). Speculation and Economic Stability. *The Review of Economic Studies*, 7(1), 1–27.

- Kang, W., Rouwenhorst, K. G., & Tang, K. (2020). A Tale of Two Premiums: The Role of Hedgers and Speculators in Commodity Futures Markets. *The Journal of Finance*, 75(1), 377–417.
- Keynes, J. M. (1930). *Treatise on money* (Vol. 2). Macmillan.
- Kuhnen, C. M., & Knutson, B. (2011). The Influence of Affect on Beliefs, Preferences, and Financial Decisions. *Journal of Financial and Quantitative Analysis*, 46(3), 605–626.
- Le Pen, Y., & Sévi, B. (2018). Futures Trading and the Excess Co-movement of Commodity Prices. *Review of Finance*, 22(1), 381–418.
- Lerner, J. S., Li, Y., Valdesolo, P., & Kassam, K. S. (2015). Emotion and Decision Making. *Annual Review of Psychology*, 66(1), 799–823.
- Lettau, M., Maggiori, M., & Weber, M. (2014). Conditional risk premia in currency markets and other asset classes. *Journal of Financial Economics*, 114(2), 197–225.
- Lewellen, J., Nagel, S., & Shanken, J. (2010). A skeptical appraisal of asset pricing tests. *Journal of Financial Economics*, 20.
- Loewenstein, G. F., Weber, E. U., Hsee, C. K., & Welch, N. (2001). Risk as feelings. *Psychological Bulletin*, 127(2), 267–286.
- Loewenstein, G. (1996). Out of Control: Visceral Influences on Behavior. *Organizational Behavior and Human Decision Processes*, 65(3), 272–292.
- Loewenstein, G. (2000). Emotions in Economic Theory and Economic Behavior. *American Economic Review*, 90(2), 426–432.
- Loewenstein, G., & Lerner, J. S. (2003). The role of affect in decision making. In *Handbook of affective sciences* (pp. 619–642). Oxford University Press.
- Ludvigson, S. C. (2004). Consumer Confidence and Consumer Spending. *Journal of Economic Perspectives*, 18(2), 29–50.
- Nagel, S. (2012). Evaporating Liquidity. *Review of Financial Studies*, 25(7), 2005–2039.
- Pham, M. T. (2007). Emotion and Rationality: A Critical Review and Interpretation of Empirical Evidence. *Review of General Psychology*, 11(2), 155–178.
- Schwarz, N. (2000). Emotion, cognition, and decision making. *Cognition & Emotion*, 14(4), 433–440.
- Shanken, J. (1992). On the Estimation of Beta-Pricing Models. *Review of Financial Studies*, 5(1), 1–33.
- Smales, L. A. (2014). News sentiment in the gold futures market. *Journal of Banking & Finance*, 49, 275–286.
- Solomon, R. C. (1993). *The passions: Emotions and the meaning of life*. Hackett Pub. Co.

- Szymanowska, M., De Roon, F., Nijman, T., & Van Den Goorbergh, R. (2014). An Anatomy of Commodity Futures Risk Premia. *The Journal of Finance*, 69(1), 453–482.
- Tang, K., & Xiong, W. (2012). Index Investment and the Financialization of Commodities. *Financial Analysts Journal*, 68(6), 54–74.
- Tetlock, P. C. (2007). Giving Content to Investor Sentiment: The Role of Media in the Stock Market. *The Journal of Finance*, 62(3), 1139–1168.
- Tetlock, P. C., Saar-Tsechansky, M., & Macskassy, S. (2008). More Than Words: Quantifying Language to Measure Firms' Fundamentals. *The Journal of Finance*, 63(3), 1437–1467.
- Vuilleumier, P. (2005). How brains beware: Neural mechanisms of emotional attention. *Trends in Cognitive Sciences*, 9(12), 585–594.
- Wahl-Jorgensen, K. (2020). An Emotional Turn in Journalism Studies? *Digital Journalism*, 8(2), 175–194.
- Zou, H. (2006). The Adaptive Lasso and Its Oracle Properties. *Journal of the American Statistical Association*, 101(476), 1418–1429.

Tables

Table 1: **Descriptive statistics of main factors**

This table reports summary statistics of the main factors used in this analysis, including news sentiment, emotion intensity, media coverage, basis, momentum, and basis-momentum. The sample covers 26 commodity futures and a period from Jan 1998 to Feb 2020.

	Mean	St. dev	Min	Max
News sentiment	-0.079	0.078	-0.725	0.309
Emotion intensity	0.317	0.121	-0.415	0.882
Media coverage	46.843	113.027	0.000	1859.617
Basis	0.009	0.029	-0.284	0.276
Momentum	0.022	0.331	-0.777	3.143
Basis-momentum	-0.016	0.062	-0.474	0.977

Table 2: **Correlation between media emotion intensity with other media factors**

This table presents the correlation between media emotion intensity and news sentiment, as well as media coverage. The table includes the correlation for the full sample and separate correlations for observations with positive news sentiment and negative news sentiment. Additionally, it provides the time-series correlation, which represents the average correlation by commodity, and the cross-section correlation, which represents the average correlation by month.

	News sentiment	Media coverage
Full sample correlation	0.099	0.160
Positive sentiment	0.177	0.113
Negative sentiment	-0.116	0.194
Time-series correlation	-0.046	0.303
Cross-section correlation	0.059	0.119

Table 3: **Performance of long-short portfolios sorted by media and benchmark factors**

This table reports the performance of the long-short portfolio sorted by media emotion intensity, news sentiment, media coverage and other benchmark factors. At the end of each month, commodity futures are sorted into three portfolios: High4 (four commodity futures with the highest values of the signal), Low4 (four commodity futures with the lowest values of the signal), and Mid (the remaining commodity futures available at the end of the month). High4 - Low4 is the portfolio of longing on the High4 portfolio and shorting on the Low4 portfolio. Each portfolio is held for a month. The nearby return of a portfolio is calculated as the equal-weighted average of the returns of the first nearby contracts in this portfolio. The spreading return is the difference between the equal-weighted average of the returns of the first and second nearby contracts. The returns are computed for the period beginning in February 1998 and ending in February 2020.

		Nearby return			Spreading return	
		High 4	Mid	Low 4	High4 - Low 4	High4 - Low 4
Emotion intensity	Ave. Ret	11.16%	0.19%	-2.35%	13.51%	3.35%
	t-stat	2.62	0.06	-0.65	2.95	4.97
	Sharpe	0.56	0.01	-0.14	0.63	1.06
Media coverage	Ave. Ret	-3.55%	1.79%	5.62%	-9.16%	-2.11%
	t-stat	-0.92	0.62	1.44	-2.69	-3.88
	Sharpe	-0.20	0.13	0.31	-0.57	-0.82
News sentiment	Ave. Ret	6.25%	0.54%	0.65%	5.60%	0.72%
	t-stat	1.37	0.18	0.15	1.13	0.76
	Sharpe	0.29	0.04	0.03	0.24	0.16
Basis	Ave. Ret	-2.57%	2.43%	2.92%	-5.49%	1.47%
	t-stat	-0.67	0.79	0.65	-1.17	1.41
	Sharpe	-0.14	0.17	0.14	-0.25	0.30
Momentum	Ave. Ret	6.21%	1.68%	-3.36%	9.57%	-1.18%
	t-stat	1.34	0.58	-0.76	1.71	-1.24
	Sharpe	0.29	0.12	-0.16	0.36	-0.26
Basis-momentum	Ave. Ret	8.22%	1.11%	-4.28%	12.50%	4.64%
	t-stat	1.91	0.36	-1.10	2.66	4.78
	Sharpe	0.41	0.08	-0.23	0.57	1.02

Table 4: **Robustness for the performance of the long-short strategy using media emotion intensity**

This table presents the robustness checks conducted to assess the performance of the long-short strategy using media emotion intensity. The second column displays the outcomes when longing six commodities with the highest factor values and shorting six commodities with the lowest factor values. The third column shows the results for longing eight commodities with the highest factor values and shorting eight commodities with the lowest factor values. The last column demonstrates the performance of the High4-Low4 portfolio, which is sorted by media emotion intensity after removing commodities that have an occurrence of more than 70% in the High4 and Low4 portfolios sorted by media emotion intensity for 26 commodities.

	High6-Low6	High8-Low8	Drop 70% High4-Low4
Nearby returns			
Average return	10.20%	7.19%	10.02%
t-stat	2.56	2.05	2.15
Sharpe ratio	0.55	0.44	0.46
Spreading returns			
Average return	2.47%	1.68%	1.56%
t-stat	3.94	3.10	2.13
Sharpe ratio	0.84	0.66	0.45

Table 5: **Time-series spanning test for the performance of long-short portfolio sorted by media emotion intensity**

Table 5 presents the results of the time-series test for the performance of the media emotion intensity long-short portfolio. The factors are labeled with "(nb)" for nearby returns and "(sp)" for spreading returns. This analysis uses three sets of factors suggested by previous literature: (1) three factors from Szymanowska et al. (2014), namely the nearby return of the High - Low portfolio sorted by basis and the spreading returns of the High and Low portfolios sorted by basis, (2) two factors from Bakshi et al. (2019), namely the nearby average return of commodity futures and the nearby return of High4 - Low4 portfolio sorted by momentum, (3) two factors from Boons and Prado (2019), namely the nearby return and spreading return of portfolios sorted by basis-momentum. The standard errors of coefficients are corrected using the Newey West method with one lag. Coefficients marked with ***, **, and * are statistically significant at the 1%, 5%, and 10% levels, respectively. The results are estimated for the sample period from February 1998 to February 2020.

MODEL	Boons and Prado (2019) and Bakshi et al. (2017)		Boons and Prado (2019) and Szymanowska et al. (2014)	
	1	2	3	4
	Nearby return	Spreading return	Nearby return	Spreading return
Alpha	0.1135**	0.0243***	0.1172**	0.0187**
t-stat (Alpha)	(2.3020)	(3.2127)	(2.2645)	(2.5460)
Basis-momentum (nb)	0.0260	-0.0064	0.0766	-0.0090
Basis momentum (sp)	-0.2682	0.2152***	-0.3289	0.1780**
Basis (nb)	-0.1103	-0.0080	-0.1730**	-0.0049
Basis (HIGH4 - sp)			-0.0848	-0.3426***
Basis (LOW4 - sp)			-0.2133	-0.1995**
Momentum (nb)	0.1044*	0.0086		
Average (nb)	0.3020***	0.0062		
Media coverage (nb)	0.0716	0.0193	0.0583	0.0146
Media coverage (sp)	-0.7817	-0.0211	-0.5703	0.0852
R-squared	0.0899	0.1096	0.0426	0.2422
GRS-F		15.0534		11.2556
p-value (GRS test)		0.0005		0.0036

Table 6: **Independent double sorting on media emotion intensity and other factors**

This table reports the nearby returns of portfolios independently double-sorted by media emotion intensity and other variables. We first report the return with t-statistics of the portfolios single-sorted by each factor in the third and fourth columns. For media emotion intensity, media coverage, momentum, and basis-momentum, we sort commodities into High and Low portfolios based on the factor's median. For basis, we sort commodities into two portfolios: contango and backwardation. All the portfolios are created by sorting commodities using factor value at the end of each month and then held for one month. The last six columns report the returns of the portfolios sorted by media emotion intensity and each variable.

		Media emotion intensity					
		Single sorting			High		
		Ave. Ret	t-stat	Ave. Ret	t-stat	Ave. Ret	t-stat
Emotion intensity	High	4.48	1.31				
	Low	-1.56	-0.51				
	High-Low	6.04	2.10				
Media coverage	High	0.73	0.23	0.12	0.03	0.11	0.03
	Low	2.42	0.80	9.17	2.34	-3.53	-1.11
	High-Low	-1.69	-0.89	-9.05	-3.14	3.64	1.17
Basis	Cont.	-0.63	-0.21	0.45	0.12	-3.41	-1.02
	Back.	7.87	1.86	11.53	2.34	-2.36	-0.48
	Cont.-Back.	-8.11	-2.26	-11.08	-2.54	-1.05	-0.22
Momentum	High	4.94	1.48	8.61	2.22	0.94	0.26
	Low	-1.94	-0.61	-1.52	-0.37	-3.92	-1.13
	High-Low	6.88	2.30	10.13	2.42	4.86	1.31
Basis-momentum	High	6.00	1.77	9.98	2.54	1.13	0.29
	Low	-3.27	-1.06	-2.10	-0.53	-4.56	-1.38
	High-Low	9.27	3.34	12.08	2.92	5.69	1.64

Table 7: Impact of media emotion intensity on trading tendency of commercial and non-commercial traders

This table presents the cross-section effects of media emotion intensity on commercial shorting pressure and non-commercial longing pressure. Commercial shorting pressure is calculated as the difference between short and long positions of commercial traders, scaled by open interest. Non-commercial longing pressure is calculated as the difference between long and short positions of non-commercial traders, scaled by open interest. This analysis utilizes the Fama Macbeth cross-section regression to evaluate the cross-sectional impact of media emotion intensity. The coefficients' standard errors are adjusted using the Newey-West method with one lag. Coefficients marked with ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

VARIABLES	Commercial shorting pressure (t)					Non-commercial longing pressure (t)				
	(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)
Emotion intensity (t-1)	0.635*** (14.926)		0.540*** (12.110)		0.517*** (10.828)	0.360*** (10.863)		0.277*** (7.880)		0.228*** (5.859)
New sentiment (t-1)		0.887*** (13.266)	0.852*** (14.007)	0.652*** (9.199)	0.706*** (10.073)		0.545*** (10.083)	0.5259*** (9.850)	0.355*** (5.940)	0.372*** (5.976)
Return (t-1)		0.416*** (8.600)	0.466*** (9.749)	0.259*** (4.330)	0.303*** (4.883)		0.456*** (11.012)	0.4916*** (11.369)	0.284*** (5.398)	0.311*** (5.573)
Basis (t-1)				-0.322* (-1.920)	-0.096 (-0.526)				-0.493*** (-3.273)	-0.425*** (-2.670)
Momentum (t-1)				0.224*** (9.065)	0.206*** (7.855)				0.246*** (10.518)	0.238*** (9.790)
Basis-momentum (t-1)				0.313*** (3.041)	0.203** (2.056)				-0.016 (-0.197)	-0.095 (-1.158)
Observations	5,636	5,636	5,636	5,477	5,477	5,636	5,636	5,636	5,477	5,477
Adj R-squared	0.096	0.119	0.194	0.156	0.223	0.049	0.097	0.136	0.159	0.192

Table 8: **Impact of media emotion intensity on trader positions with the moderation of media coverage**

This table reports the results of regressions of commercial shorting pressure and non-commercial longing pressure on media emotion intensity and the interaction between media emotion intensity and the dummy variable for low media coverage. The low media coverage is measured cross-sectionally and takes the value of 1 if a commodity has media coverage below the median value of all commodities in each month. This analysis utilizes the Fama Macbeth cross-section regression. In the cross-section test, the standard errors of coefficients are corrected using Newey West with one lag.

VARIABLES	Commercial shorting pressure (t)			Non-commercial longing pressure (t)		
	(1)	(2)	(3)	(1)	(2)	(3)
Emotion intensity (t-1)	0.412*** (10.973)	0.346*** (8.795)	0.355*** (8.251)	0.204*** (6.517)	0.141*** (4.240)	0.131*** (3.515)
Emotion intensity (t-1) × Media coverage (t-1)	0.343*** (15.915)	0.315*** (15.071)	0.283*** (11.452)	0.237*** (14.887)	0.216*** (13.754)	0.166*** (8.711)
New sentiment (t-1)		0.769*** (12.207)	0.613*** (8.325)		0.459*** (7.942)	0.313*** (4.617)
Return (t-1)		0.477*** (10.155)	0.359*** (5.920)		0.501*** (11.352)	0.354*** (6.245)
Basis (t-1)			-0.087 (-0.477)			-0.439*** (-2.710)
Momentum (t-1)			0.191*** (7.248)			0.225*** (9.059)
Basis-momentum (t-1)			0.136 (1.334)			-0.132 (-1.529)
Observations	5,636	5,636	5,477	5,636	5,636	5,477
Adj R-squared	0.223	0.381	0.521	0.165	0.323	0.490

Table 9: Media emotion intensity and the cross-section of commodity portfolio returns

This table reports the cross-section tests for media emotion intensity and other benchmark factors for the sample from February 1998 to February 2020 at the portfolio level. Model (1) presents the test for the factors introduced by Bakshi et al. (2019) and Boons and Prado (2019). Models (2) and (3) test whether the risk of media emotion intensity nearby and spreading returns are priced in the cross-section of portfolio returns against the factors in Model (1). Model (4) presents the test for the factors introduced by Szymanowska et al. (2014) and Boons and Prado (2019). Models (5) and (6) test whether the risk of media emotion intensity nearby and spreading returns are priced in the cross-section of portfolio returns against the factors in Model (4). This analysis used 42 portfolios: including (1) 30 portfolios by sorting each factor of basis, momentum, basis-momentum, media emotion intensity, and media coverage into three groups (High4, Mid, Low4) and separating to nearby and spreading strategies; (2) 12 portfolios by the two above strategies for commodity types (Energy, Grains, Livestock, Metals, Softs and Food oils). Then, this study conducts the cross-section regression of the average returns of portfolios against their exposures to factors. The estimated risk premium of each factor is reported with the Newey West t-statistic from Fama-Macbeth estimation in the parentheses underneath and the t-statistic after correcting standard error using Shanken (1992) method in the square brackets.

MODEL	Portfolio return					
	Boons and Prado (2019) and Bakshi et al. (2017)		Boons and Prado (2019) and Szymanowska et al. (2014)			
	1	2	3	4	5	6
Basis (nb)	-0.007* (-1.727) [-1.258]	-0.005 (-1.233) [-0.877]	-0.005 (-1.241) [-0.880]	-0.011** (-2.402) [-1.757]	-0.007 (-1.600) [-1.146]	-0.007* (-1.665) [-1.165]
Basis (High4 - sp)				0.000 (0.092) [0.080]	-0.000 (-0.190) [-0.165]	-0.000 (-0.199) [-0.171]
Basis (Low4 - sp)				-0.004** (-2.490) [-2.039]	-0.002 (-1.628) [-1.316]	-0.002 (-1.509) [-1.233]

Average	0.003 (1.145)	0.003 (1.029)	0.003 (1.042)
Momentum (nb)	[0.845]	[0.747]	[0.759]
	0.011**	0.008*	0.008*
	(2.294)	(1.688)	(1.690)
	[1.557]	[1.138]	[1.139]
Basis-momentum (nb)	0.010**	0.010**	0.012**
	(2.378)	(2.401)	(2.355)
	[1.748]	[1.735]	[1.745]
Basis-momentum (sp)	-0.000	0.003***	0.001
	(-0.081)	(2.791)	(0.673)
	[-0.072]	[2.253]	[0.595]
Emotion intensity (nb)	0.012***	0.012***	0.013***
	(2.938)	(2.930)	(2.953)
Emotion intensity (sp)	[2.106]	[2.098]	[2.073]
	0.001	0.001	0.002
	(0.926)	(0.926)	(1.398)
	[0.761]	[0.761]	[1.107]
Constant	-0.001***	-0.001***	-0.001
	(-5.218)	(-4.333)	(-1.337)
	[-5.110]	[-4.241]	[-1.051]
Observation	11,130	11,130	11,130
R-squared	0.517	0.613	0.527
		0.430	0.601

Table 10: Media emotion intensity and the cross-section of individual commodity returns

This table reports the cross-section tests for media emotion intensity and other benchmark factors for the sample from February 1998 to February 2020 at the individual commodity level. Model (1) presents the test for the factors introduced by Bakshi et al. (2019) and Boons and Prado (2019). Model (2) and (3) tests whether the risk of media emotion intensity nearby and spreading returns are priced in the cross-section of portfolio returns against the factors in Model (1). Model (4) presents the test for the factors introduced by Szymanowska et al. (2014) and Boons and Prado (2019). Model (5) and (6) tests whether the risk of media emotion intensity nearby and spreading returns are priced in the cross-section of portfolio returns against the factors in Model (4). To run the cross-section tests at the commodity level, this analysis employs Fama-Macbeth (1973) regression for the nearby and spreading returns of 26 commodity futures. In the first stage, this study runs the rolling time-series regression with a one-year rolling of daily returns. In the second stage, this study runs the monthly cross-section regression of nearby and spreading returns of commodity futures on their time-varying exposures to factors. The table reports the estimated risk premiums of factors with their t-statistic corrected using the Newey West method for one lag in the parenthesis and their t-statistic corrected using Shanken (1992) method in the square brackets.

MODEL	Individual commodity futures return					
	Boons and Prado (2019) and Bakshi et al. (2017)			Boons and Prado (2019) and Szymanowska et al. (2014)		
	1	2	3	4	5	6
Basis (nb)	-0.007 (-1.140) [-0.976]	-0.005 (-0.821) [-0.686]	-0.007 (-1.058) [-0.887]	-0.006 (-0.928) [-0.747]	-0.007 (-1.148) [-0.908]	-0.008 (-1.341) [-1.018]
Basis (High4 - sp)				0.002 (1.407) [1.236]	0.001 (0.862) [0.767]	0.001 (0.954) [0.799]
Basis (Low4 - sp)				0.001 (0.470) [0.405]	0.002 (1.014) [0.857]	0.002 (1.203) [0.939]

Average	0.004 (1.419)	0.003 (1.367)	0.003 (1.372)
Momentum (nb)	[1.010]	[0.977]	[0.980]
	0.002 (0.353)	0.006 (0.921)	0.007 (1.124)
	[0.273]	[0.722]	[0.878]
Basis-momentum (nb)	0.015***	0.016***	0.015***
	(2.770)	(2.840)	(2.768)
	[2.179]	[2.280]	[2.221]
Basis-momentum (sp)	0.003 (1.500)	0.002 (1.484)	0.002 (1.236)
	[1.348]	[1.286]	[1.090]
Emotion intensity (nb)		0.011**	0.011**
		(2.295)	(2.191)
		[1.760]	[1.677]
Emotion intensity (sp)			0.001 (1.153)
			[0.953]
Constant	-0.001*** (-3.467)	-0.001*** (-3.755)	-0.001*** (-3.586)
	[-3.386]	[-3.600]	[-3.439]
			[-1.044]
Observation	12,083	12,083	12,083
R-squared	0.412	0.488	0.521
			0.368
			0.438
			0.474

Table 11: **LASSO regression for selecting relevant factors**

This table summarizes the relative frequency of being selected for each factor in the first step of the two-pass regression with double-selection LASSO introduced by Feng et al. (2020). This test employs LASSO cross-section regression to regress portfolios' average returns on the covariances between the portfolio returns with each factor. This technique will select the factors that are the most relevant to the cross-section of portfolio returns. This study employs three types of LASSO models: standard LASSO, adaptive LASSO introduced by Zou (2006) and plug-in LASSO from Belloni et al. (2012). First, the analysis generates 200 sub-samples from our main sample. Each sub-sample contains 75% observations of the whole sample. Second, the test runs the LASSO model on the 200 samples and counts the number of times each factor is selected among these 200 samples. The numbers reported in this table are the relative frequencies of each factor that appears in the set of selected factors.

	LASSO technique		
	Standard	Adaptive	Plugin
Emotion intensity (nb)	100.0%	100.0%	87.5%
Emotion intensity (sp)	80.0%	34.5%	31.5%
Average	57.0%	28.5%	21.5%
Basis (nb)	49.0%	23.0%	33.0%
Basis (High4 - sp)	21.5%	12.0%	6.0%
Basis (Low4 - sp)	27.0%	13.5%	0.0%
Momentum (nb)	53.5%	28.5%	47.5%
Basis-momentum (nb)	72.0%	53.5%	82.0%
Basis-momentum (sp)	62.5%	45.5%	7.0%
Media coverage (nb)	79.0%	50.0%	1.0%
Media coverage (sp)	38.0%	15.0%	0.0%

Internet appendix

A.1 - Thomson Reuters MarketPsych indices data and lexical analysis

This study uses the Thomson Reuters MarketPsych Indices (TRMI) to calculate media emotion intensity. TRMI is built on a system of indicators that reflect a variety of psychological aspects extracted from media news and analyzed by the Thomson Reuters news reading system. The strength of TRMI is that this index system uses a unique natural language processing algorithm built on a complicated system of lexicon dedicated to business and investment, and advanced grammatical analysis for news and social media textual analysis.

Traditionally, textual analysis often employs the “bag of words” approach. This method defines a dictionary of sentiment words and scores each word and phrase in a text using this predefined dictionary. There are some potential challenges with sentimental textual analysis. First, polarized measures such as sentiment with two aspects (positive and negative) might be confusing to a certain extent. Some information might be positive for some investors but negative for others, depending on the context. Second, the dictionary is required to be updated frequently to fit the current context. Third, the content should be carefully analyzed to focus only on the related information to a particular asset. Thomson Reuters overcomes these challenges by creating

dynamic dictionaries for each aspect of information and regularly updating them to suit the financial and business environment. To associate the text's content with a particular asset, the system maintains a list of asset names and their aliases. This list is periodically updated and human-reviewed to maintain its accuracy. To exclude unrelated tokens generated by keyword scanning, the system uses a supervised machine learning algorithm to determine which tokens correspond with the asset names and which do not. For instance, when associating news material with gold, the machine can scan the token "gold medals" and relate the phrase to the Olympic Games rather than the gold commodity. As in the preceding example, the algorithm will employ anti-correlation filtering and case sensitivity to eliminate the perplexing tokens. Additionally, the system uses a correlation filter to determine whether the contents are associated with the asset. For instance, if the algorithm encounters the content "I love eating corn," this content will not be counted for "corn". The algorithm counts references only if they contain important identifiers such as "return" or "futures".

The Thomson Reuters news reading system collects words and phrases and categorizes them into two groups: psychological and factual. The psychological category encompasses words and phrases such as joy, hope, fear, stress, happiness, hate, and love that reflect psychological aspects. The factual category contains factual information about a subject, such as earnings, price, or return. The system then examines the attributes of each word. TRMI calls words or phrases with their attributes are "Variable". Accounting, earning,

ambiguity, and anxiety are all examples of words. In these examples, accounting and earning might connect to factual information, while ambiguity and fear are emotional aspects. When scanning the text, the system will add the grammatical tense to the words and phrases and add the directional sentiment to them (if any) to form Variables. For instance, a sentence contains good information on accounting, and this information is from the past so that the system will save it as a Variable, namely “AccountingGood_p”.

With regard to scoring a Variable, this study will focus on an example of a sentence: “Experts expect gold will have much higher price next month”. First, the system will scan the word “gold” to associate this information with the commodity “gold”. Second, they identify the word “price” as a word in the dictionary. Third, the word “expect” relates the information to the future tense. Fourth, identify the word “higher” as an “up” word. Hence, the system will record a Variable “PriceUp_f” to refer to a piece of information on higher prices in the future for the asset “gold”. A standard score for a Variable is 1. However, because the word “higher” accompanies the word “much”, which intensifies the level of “higher”, the score for the Variable “PriceUp_f” in this context is doubled to 2. If the system scans the text and finds some weaker modifiers, such as minimally, it will halve the score of the Variable to 0.5. Similarly, we can observe an example of emotional content in the sentence “I love investing in gold futures”. This sentence shows the view of the authors, which is not factual information of the asset “gold”. In this example, the word “love” refers to the psychological aspect of “love”.

The word “love” also becomes a Variable “love” and receives a score of 1. This Variable is recorded for the asset “gold” and the category “love”.

A.2 - News sentiment measure

The term “news sentiment” is frequently used to refer to two sentimental aspects of media news, positive and negative. García (2013), Tetlock (2007), and Tetlock et al. (2008) measure positive and negative news sentiment by the proportion of positive and negative words in the news articles based on predefined dictionaries suggested in the literature. This study uses the sentiment indicator in TRMI data to measure news sentiment for each commodity. TRMI uses the self-defined dictionary for positive and negative words to calculate news sentiment. This dictionary is designed specifically for investment and business contexts and is frequently updated and validated by humans. For a specific period, define the function $I_{sen}(v)$ as

$$I_{sen}(t, v) = \begin{cases} +1 & \text{if } v \text{ conveys a positive view on the asset} \\ -1 & \text{if } v \text{ conveys a negative view on the asset.} \end{cases} \quad (11)$$

The news sentiment of the asset a for the specific period is calculated as

$$News\ sentiment(a) = \frac{\sum_{v \in V(a)} (I_{sen}(v) \times S(v))}{buzz(a)}. \quad (12)$$

This sentiment indicator is the net of positive to negative news sentiment and ranges from -1 to 1. A positive value reflects the positive average tone

of media toward the asset. In comparison, a negative value indicates that the media, on average, views the asset negatively. Many previous studies, such as García (2013) and Smales (2014), measure news sentiment for each news item and then aggregate the news-level sentiment to calculate the daily news sentiment. This study aggregates the monthly news sentiment from word-level sentiments. This measure is to consider the length of the news content in each news item. Further, this study features long-short strategies to evaluate the factors. Separating positive from negative sentiments might not accurately reflect the media's overall view of an asset. Therefore, the net news sentiment is preferred to rank the commodity futures and test the performance of the portfolio constructed based on this factor.

A3 - Discussion on whether media emotion intensity premium is subsumed by other risk sources

A3.1 - Hedging pressure

The role of hedging pressure stretches back to the view of Keynes (1930) on the theory of normal backwardation. According to this theory, hedgers participate in the futures market to hedge against spot price fluctuations. As hedgers take more short positions in the market, speculators share the risk with hedgers by setting the futures price lower than the expected spot price. In this way, the futures price is affected by hedging pressure. If the media emotion intensity implies hedging pressure, we can observe the performance of the media emotion intensity factor being adjusted in response to hedging

pressure.

To measure hedging pressure, we employ the data on trader positions from the Commodity Futures Trading Commission (CFTC). The CFTC publishes weekly reports, Commitment of Trader, that summarize the commercial and non-commercial traders' long and short positions in the commodity futures market. These reports are frequently provided on Friday of each week and contain data for the preceding Tuesday-to-Tuesday period. There is no perfect method for aggregating weekly data to monthly data. In this study, we assign a week (Tuesday to Tuesday) to a specific month if the last day of this week falls on that month. Kang et al. (2020) smooth the hedging pressure measure by taking the 52-week moving average of the weekly hedging pressure and show that this measure better captures the true hedging pressure in the commodity futures market. We use the same rolling period to calculate our hedging pressure variable (*12-month HP*) but modify it to suit the monthly frequency (12-month rolling). Furthermore, we also consider short-term HP (*monthly HP*).

First, we calculate the time series and cross-section correlations between hedging pressure and other factors to determine whether it interacts with our media emotion intensity factor and other benchmarks. We employ the Boons and Prado (2019) approach. To calculate the time-series correlation, we first calculate the correlation between *HP* and the factor *F* for each commodity and then take the median of the correlations (across all commodities) as the time-series correlation between *HP* and the factor *F*. By comparison, the

cross-section correlation between HP and the factor F is determined as the median of a series of monthly cross-section correlations between HP and F . These numbers are shown in Table A7.1 Panel A. Only momentum is found to be highly linked with hedging pressure over time. While it is clear that basis-momentum and basis have a minor correlation with hedging pressure, media emotion intensity is largely unrelated. Regarding cross-section associations, media emotion intensity has the strongest link with hedging pressure, at 0.37, followed by momentum (0.30). This could suggest that media emotion intensity might be partly related to hedging pressure in the commodity futures market.

[Table A7.1 here]

Further, we examine whether the long-short portfolio's abnormal return is primarily explained by hedging pressure. Panel B of Table A7.1 summarized the findings of the spanning test for the media emotion intensity long-short portfolio return when compared with hedging pressure, other benchmarks, and all of them. We begin with sorting commodities into three portfolios (High4, Mid and Low4) based on hedging pressure for this spanning test. The nearby return of the High4-Low4 portfolio sorted by hedging pressure will be used to test the association of this factor to the media emotion intensity factor. The standard errors of all coefficients are adjusted using the Newey West method with one lag. Further, we compare media emotion intensity and basis-momentum performance using the same spanning test for

the basis-momentum long-short portfolio. Model 1 shows that the return of hedging pressure portfolio is significantly correlated with the return of media emotion intensity portfolio; even after controlling for other factors, the correlation between hedging pressure and media emotion intensity remains highly significant. Model 2 indicates that media emotion intensity can generate a high abnormal return of about 12% annualized, even after controlling for the main benchmarks. When hedging pressure is considered, the alpha decreases to 7.74 %, and the corresponding significance decreases from $t\text{-stat} = 2.56$ to $t\text{-stat} = 1.83$. This result implies that media emotion intensity abnormal return is linked to hedging pressure. Media emotion intensity still generates a high abnormal return of nearly 8% after controlling for hedging pressure, implying that a substantial portion of the media emotion intensity return cannot be explained by hedging pressure or other benchmarks. When controlling for media emotion intensity and other benchmarks, basis-momentum can generate a 10% abnormal return. However, adding hedging pressure to the test reduces the alpha of basis-momentum to 7.85%, and the $t\text{-stat}$ also decreases significantly from 2.25 to 1.80. Furthermore, we discover a substantial link between the long-short return of the hedging pressure portfolio and the long-short return of the basis-momentum portfolio. This result also suggests that the basis-momentum abnormal return is strongly related to hedging pressure. After accounting for hedging pressure, both media emotion intensity and basis-momentum continue to perform well and generate large abnormal returns at a similar power and significance level.

Although the spanning test indicates that a portion of the media emotion intensity's abnormal return might be attributed to hedging pressure, the cross-section analysis should be used to examine whether the media emotion intensity premium still significantly predicts the cross-section of commodity futures returns if hedging pressure is controlled. The result of this cross-section test is shown in Pane C of Table A7.1. We again assess the relevance of factors using Fama-Macbeth regression and Shanken (1992) correction. Model (1) includes the average factor representing market risk and hedging pressure. The result shows that the premium of hedging pressure is high and significant at 1.7% a month (nearly 20% annualized). When adding media emotion intensity to form Model (3), the risk price of hedging pressure falls to 1.2% a month (t-stat decreases from 2.58 to 1.77, suggesting that media emotion intensity does relate to hedging pressure. First, the positive and significant premium of hedging pressure is consistent with the theory of backwardation. Also, this estimation is consistent with Kang et al. (2020) that long-term risk is mainly driven by hedging pressure from commercial traders. Second, the decline in the hedging pressure premium while the premium of media emotion intensity remains unchanged indicates that media emotion intensity might partly link to hedging pressure but not be subsumed by this risk factor.

Interestingly, when hedging pressure and basis-momentum are combined in the same model, the risk price of basis-momentum falls to 0.8% per month and loses its significance. This implies that hedging pressure risk is also

related and explains basis-momentum risk. When all factors are combined in the same test, we notice that the premium of hedging pressure decreases dramatically to 0.7% per month and becomes insignificant. We further observe that, in this model, the premium of basis-momentum changes a little, but the momentum factor substantially loses its significance and premium compared to Table 11. In summary, the cross-section test results indicate that the risk of hedging pressure might be largely linked to media emotion intensity and momentum. Unlike momentum, the estimated media emotion intensity premiums are consistent and high across all models. This conclusion is important because it indicates that a substantial portion of media emotion intensity premium cannot be explained by other factors.

A3.2 - Inventory

The evidence on the relation between media emotion intensity risk and hedging pressure suggests a link between media emotion intensity and the inventory state of commodities. Gorton et al. (2013) suggest that the physical inventory level is what drives the risk premiums of commodity futures. They note that basis and past futures return could capture the inventory level of commodities and explain the risk premiums. The origins of these interactions can be traced back to storage theory (Kaldor, 1939). At the high inventory level, the commodity supply is high and the futures price pattern is more likely to be contango. This happens because the futures price will increase to compensate for the storage cost. In contrast, when a commodity's inventory

is low, the spot price will increase to reflect the scarcity of supplies. Therefore, the futures price pattern is more likely to be backwardation. According to Gorton et al. (2013), a low inventory level is frequently related to a negative basis (backwardation) and higher past futures returns. Thus, inventory risk can be represented by the basis and momentum of commodity futures returns.

The single sorts in Table 3 and Table 6 already demonstrate that backwardation and high momentum portfolios have a larger return than their counterparts. This result is consistent with Boons and Prado (2019) and Gorton et al. (2013), and supports the theory of storage. However, the spanning tests in Table 5 and Table A7.1 reveal that when controlling for basis, momentum, and hedging pressure, the alphas are still high, above 7.6%, suggesting that these proxies for the inventory level cannot explain a large proportion of media emotion intensity abnormal returns. The cross-section test in Table 11 and Table A3 also confirm that the media emotion intensity premium is not subsumed by hedging pressure, basis, and momentum factors. Unfortunately, we do not have the direct measures for inventory and storage level as in Boons and Prado (2019). However, these results might suggest that a large part of media emotion intensity premium is not explained by the inventory level of commodities.

A3.3 - Market volatility and average volatility

According to Boons and Prado (2019), market volatility risk is priced in the cross-section of commodity futures returns. The concept of market

liquidity stretches all the way back to Brunnermeier and Pedersen (2009). They contended that market liquidity is related to market volatility and that traders' ability to provide liquidity relies on their funding availability. Additionally, they demonstrated that speculators' funding availability might explain risk premiums. Boons and Prado (2019) argued that the negative risk premium associated with market volatility indicates that investors are willing to pay a premium to offset the increased volatility. In accordance with Boons and Prado (2019), we additionally examine whether media emotion intensity links to market volatility risk.

This study uses Boons and Prado (2019) methodology to calculate market and average volatility. We begin by building a market portfolio that is equally weighted in commodity futures to calculate market volatility. Then, the monthly market volatility is determined as the sum square of the market portfolio's daily returns. The average volatility is calculated as the mean of the monthly volatility of commodity futures' daily returns, where the monthly volatility of each commodity futures equals the sum square of its daily returns. First, we examine whether the market and average volatility can be used to predict the returns of portfolios sorted by media emotion intensity. The outcome is shown in Table A7.2. To conduct the test, we estimate WLS regression models of the nearby returns of the High4, Low4, Mid, and High4-Low4 portfolios at the month $t + 1$ against the market volatility and the average volatility at the time t . These regressions are similar to those described in Boons and Prado (2019). In WLS regression, the weight

is the inverted conditional volatility, measured by the standard deviation of the last 12 months' returns. The result indicates that both market volatility and average volatility have limited predictive potential, with R^2 values of approximately 1% for the High4, Low4, and Mid portfolios and even as low as 0.01% for the High4-Low4 portfolio. Additionally, the t-stat for the market and average volatility coefficients in the High4-Low4 return model is quite minimal, at roughly 0.2. This indicates that market volatility or average volatility only explains a small portion of the following month performance of the long-short portfolio ranked by media emotion intensity.

[Table A7.2 here]

Furthermore, we investigate whether the media emotion intensity premium is subsumed by market volatility and average volatility. We use Fama-Macbeth regression with Shanken (1992) adjustment to do the cross-section test. The results of this test are summarized in Table A7.3. We begin with only the market factor and each volatility factor and then add the media emotion intensity factor into the model, followed by other benchmark factors. Model (1) indicates that the risk price of market volatility is -7.2% per month when only the average factor is used. This estimate is comparable to the -8.0% monthly recorded by Boons and Prado (2019). Additionally, our estimate of -2.8% in Model (2) is consistent with the 2019 paper's estimate of -2.4%. The only difference is that market volatility and average volatility are not significantly priced in the cross-section in our test. However, this

may be because our sample size is smaller than that of Boons and Prado (2019). The Shanken t-statistic for market volatility is -1.42, which is close to being significant. The Shanken t-stat for average volatility, on the other hand, is -0.34, which is far from being significant. Interestingly, the media emotion intensity factor's premium remains highly constant and significant across models. The premium associated with media emotion intensity is approximately 1.2% a month (with t-stat ≈ 2), consistent with the results of all previous cross-section tests in this document. Therefore, we can conclude that the media emotion intensity premium is not subsumed by market volatility and average volatility risk.

[Table A7.3 here]

A3.4 - Market liquidity and funding liquidity

Brunnermeier and Pedersen (2009) argue that market liquidity and funding liquidity are connected to market volatility. They also point out that these two types of liquidity might link together. When the funding capital is tighter, investors appear to be less interested in trading. As a result, the market becomes increasingly illiquid, resulting in an increase in market volatility. The risk price of the market volatility measured for commodity futures markets might also suggest media emotion intensity might partly capture the market liquidity risk in the commodity futures market. We use the Ted spread as a proxy for funding illiquidity in the same way as Boons and Prado (2019). Previous studies often favor VXO over VIX owing to the

rich historical data. Also, VIX and VXO are highly correlated. In our study, we employ VIX as a proxy for equity market volatility as our sample covers the period from 1998 when VIX was already available. As increased volatility may result from the increased market and funding illiquidity, VIX also reflects the risk associated with these two types of illiquidity.

We apply the two-stage Fama Macbeth regression with Shanken (1992) correction to evaluate these liquidity proxies. We use TED spread directly in the models as TED spread does not contain a time trend component and is stationary. However, because VIX is highly volatile and non-stationary, we conduct the test using the first-order difference of VIX. This transformation is required to assist the time series regression in the first stage. First, we run the cross-section test with only the average factor and each illiquidity proxy in the first two models. Models (3) and (4) are built by supplementing Models (1) and (2) with media emotion intensity factors, respectively. Additionally, we incorporate basis-momentum, basis, and momentum factors to create Models (5) and (6). The results are reported in Table A7.4.

[Table A7.4 here]

Models (1) and (2) demonstrate that when only the average factor is controlled, both illiquidity proxies can explain the cross-section of portfolio returns. When media emotion intensity factors are included in Model (3), the estimated risk coefficient decreases from 0.227 to 0.181, and the t-stat drops to 1.365. This finding implies that media emotion intensity may

partially represent the funding liquidity risk. When the basis-momentum, basis, and momentum factors are added to Model (5), the coefficient of Ted spread continues to decrease and loses its significance. This change is consistent with Boons and Prado (2019) finding that basis-momentum also reflects funding liquidity risk. Similarly, the coefficient of the VIX factor is significant in Model (2) but less so in Model (4), highlighting that media emotion intensity might be related to funding and market liquidity. We calculate the correlation between the returns of the High4-Low4 portfolio ranked by media emotion intensity and the short-term reversal factor (Nagel, 2012) published on Kenneth French's website, following Boons and Prado (2019). A correlation of less than 0.1 indicates a weak relation between media emotion intensity premium in the commodity markets and liquidity provision in the equity market. The premium of media emotion intensity is relatively stable and consistent, at around significantly 1.2% a month. In other words, the premium of media emotion intensity is not subsumed by liquidity risks.

A3.5 - Equity market risk and Downside equity market risk

Boons and Prado (2019) suggests that both market risk and downside market risk of the equity market are priced in the cross-section of portfolio returns in the commodity futures markets. They argued that the bear equity market might discourage speculators from clearing the commodity futures market. Table A7.5 illustrates the cross-section test results when the equity market risk and equity downside market risk are considered. Similar to

other cross-section tests, we apply the two-stage Fama-Macbeth regression with Shanken (1992) correction. We use the value-weighted CRSP portfolio's excess return to represent the equity market return. If the market return is less than the average minus one standard deviation for the entire sample, the market is deemed to be on the downside. This measure is suggested by Lettau et al. (2014).

[Table A7.5 here]

Model (1) shows that equity market risk accounts for 3.1% of the cross-section of portfolio returns in the commodity futures market each month. When downside market risk is added in Model (2), the equity market and downside equity market risk premiums are 3.7% (t-stat = 1.60) and 2.7% (t-stat = 1.62), respectively. After incorporating the media emotion intensity factor into the testing models, the equity market's risk premium is decreased to 2% and is no longer significant (t-stat = 1.20). Similarly, the risk price of the downside equity market also drops to 1.5% (t-stat = 1.4). This finding implies that the media emotion intensity may relate to both the equity market risk and the equity market's downside risk. The premium associated with media emotion intensity is fairly stable at 1.2% per month and significant, as seen in the previous test. In our sample, the risk price of the downside market risk is smaller than in the sample of Boons and Prado (2019), suggesting that the commodity futures market is less affected by the sentiment from the equity market. This test also indicates that the premium of media emotion

intensity is not subsumed by equity market risk and downside equity market risk.

A3.6 - Macroeconomic factors

Given the fact that the situation of the economy has an impact on the supply and demand of commodities, the macroeconomic factor is inextricably linked to the commodity market. Due to the unique nature of different commodities, each commodity might respond differently to macroeconomic shocks. Numerous studies have established the macroeconomic influences on the returns on commodity prices (see Le Pen and Sévi, 2018). This section tests the predictive role of macroeconomic risks in the cross-section of commodity futures returns and whether the macroeconomic risks subsume the premium of media emotion intensity.

We collect ten macroeconomic variables: CPI, CPI for commodity, real M1 supply, real M2 supply, PPI for final demand of commodities, PPI for commodities in the US, PPI for manufacturing, Unemployment, US Dollar index, and Michigan consumer sentiment index. These are the key macroeconomic variables that reflect various facets of the US economy. We are aware that while several of these variables are highly correlated, they are not identical. As a result, we begin by doing a principal component analysis of these macro variables to identify the key factors containing the primary information. Because none of these variables is stationary, they are transformed to the first-order difference form to facilitate the time-series regression in the

first stage of the cross-section test. The results of the principal component analysis are summarized in Table A7.6. Panel A demonstrates that the third factor's eigenvalue is 1.09, whereas the fourth factor's eigenvalue is 0.84. As a result, following PCA analysis, we retain only three factors. Panel B reports the loading of each factor on the macroeconomic variables. Factor 1 is mainly associated with CPI, PPI and M2 supply. These factors reflect the price level of the economy. Factor 2 mostly captures the change in consumer sentiment and unemployment. Both of these variables are related to the labor market and consumer confidence (see Ludvigson, 2004). Factor 3 is tied to the M1 supply and the US dollar index, which measures the US currency's strength.

[Table A7.6 here]

The cross-section test for macroeconomic factors using two-stage Fama-Macbeth regression with Shanken, 1992 correction is summarized in Table A7.7. Only the factor related to consumer confidence and the labor market can significantly explain the cross-section of portfolio returns in our sample. Factor 1 risk is not significantly priced ($t\text{-stat} = -1.11$), showing that the inflation risk contributes only a negligible amount to the cross-section explanation. This finding contrasts with that of Szymanowska et al. (2014), who discovered that inflation risk has a larger predictive power. It should be emphasized, however, that the Shanken correction significantly reduces the t -statistic compared to the OLS estimate or Newey West correction. Furthermore, given our short sample period, it is more difficult for factors to

survive the significant thresholds. However, we can observe the relation between media emotion intensity and macroeconomic risks through the change in coefficients and their corresponding t-statistic. Feng et al. (2020) argued that we should still bring back the variables that correlate with the main factors to reduce the estimation bias for premiums.

[Table A7.7 here]

Two major findings are revealed in Table A7.7. First, when adding media emotion intensity to the model of macroeconomic factors with the average factor, the estimated coefficients of the macroeconomic factors change greatly. Also, their t-statistics decrease considerably. Particularly for Factor 2, the estimated coefficient loses its significance markedly from $t\text{-stat} = -1.80$ to $t\text{-stat} = -0.196$. This means that the premium associated with media emotion intensity is related to the risk associated with the labor market and consumer spending. Even this relationship is obscure in the previous literature. Second, after controlling for all the macroeconomic factors and other benchmarks, the estimated premium of the media emotion intensity factor remains significant and rather steady, staying at around 1.2% monthly. This result confirms that the media emotion intensity premium is not subsumed by macroeconomic risks.

A.4 - Appendix tables

Table A1: **Descriptive statistics of the 1st nearby and spreading returns of commodity futures**

This table presents the descriptive statistics of the 1st nearby and spreading returns of commodity futures. The 1st nearby contract refers to the contract closest to maturity that will not expire within the next two months. The spreading return is the difference in returns between the 1st nearby and 2nd nearby contracts. The 2nd nearby contract is the contract closest to maturity following the 1st nearby contract.

Commodity	Ticker	1st nearby return		Spreading return	
		Average	St. dev.	Average	St. dev.
Ethanol	AK	17.54%	30.50%	5.11%	7.75%
Aluminum	AL	-4.36%	14.53%	0.83%	1.20%
Soybean Oil	BO	-2.01%	25.17%	-0.97%	1.50%
Corn	C	-6.89%	26.65%	-3.03%	3.26%
Cocoa	CC	2.16%	29.87%	-0.79%	2.62%
Crude Oil	CL	6.87%	31.27%	-1.26%	2.48%
Cotton	CT	-4.82%	27.60%	-3.46%	5.30%
Palm Oil	CU	-4.98%	21.64%	0.10%	2.58%
Cattle	FC	1.01%	15.69%	-2.40%	2.88%
Gold	GC	6.60%	16.49%	-0.05%	0.26%
Copper	HG	7.82%	25.93%	-0.17%	1.07%
Heating Oil	HO	6.83%	31.27%	-0.54%	3.22%
Coffee	KC	-8.75%	31.50%	-1.18%	2.38%
Hogs	LH	-9.55%	26.34%	-8.68%	9.80%
Natural Gas	NG	-10.77%	42.87%	-4.49%	9.91%
Orange Juice	OJ	-2.08%	29.71%	-0.24%	3.45%
Palladium	PA	17.81%	35.50%	0.74%	3.77%
Platinum	PL	7.05%	21.99%	0.77%	1.84%
Brent Crude	QA	8.88%	30.64%	-0.53%	2.08%
Gasoline	RB	14.22%	33.90%	1.06%	5.98%
Rough Rice	RR	-9.47%	23.81%	-3.22%	5.36%
Canola	RS	-2.26%	20.51%	-2.03%	3.40%
Soybeans	S	4.69%	24.88%	-0.47%	2.48%
Sugar	SB	0.82%	31.93%	-1.27%	7.41%
Silver	SI	6.27%	29.40%	-0.31%	0.46%
Wheat	W	-8.56%	28.98%	-2.84%	3.36%

Table A2: **Descriptive statistics of the media factors for each commodity**

This table displays the descriptive statistics of three media factors, including media emotion intensity, media coverage, and news sentiment for each commodity. The sample covers the period from January 1998 to February 2020.

Commodity	Ticker	Emotion intensity		Media coverage		News sentiment	
		Mean	St. dev	Mean	St. dev	Mean	St. dev
Ethanol	AK	0.391	0.058	0.955	0.548	-0.014	0.059
Aluminum	AL	0.422	0.045	3.803	3.192	-0.055	0.062
Soybean oil	BO	0.232	0.172	0.064	0.065	-0.068	0.124
Corn	C	0.196	0.048	3.704	2.538	-0.066	0.035
Cocoa	CC	0.311	0.084	1.762	0.862	-0.041	0.055
Crude oil	CL	0.378	0.040	41.985	34.210	-0.118	0.041
Cotton	CT	0.263	0.070	1.804	1.097	-0.058	0.035
Palm oil	CU	0.354	0.043	1.570	0.673	-0.096	0.053
Cattle	FC	0.309	0.086	3.940	2.540	-0.128	0.036
Gold	GC	0.413	0.036	13.839	11.494	-0.037	0.048
Copper	HG	0.335	0.075	5.178	3.639	-0.082	0.055
Heating oil	HO	0.312	0.097	0.526	0.501	-0.144	0.093
Coffee	KC	0.342	0.086	2.891	1.440	-0.036	0.050
Hogs	LH	0.166	0.066	0.749	0.311	-0.127	0.044
Natural gas	NG	0.366	0.036	8.052	6.575	-0.080	0.049
Orange juice	OJ	0.263	0.139	0.208	0.171	-0.103	0.069
Palladium	PA	0.457	0.123	0.404	0.323	-0.071	0.144
Platinum	PL	0.446	0.095	1.251	1.444	-0.011	0.094
Brent crude	QA	0.425	0.112	1.127	1.072	-0.155	0.101
Gasoline	RB	0.294	0.047	5.864	4.510	-0.111	0.048
Rough rice	RR	0.313	0.102	3.388	2.937	-0.068	0.036
Canola	RS	0.203	0.084	0.459	0.243	-0.110	0.063
Soybean	S	0.186	0.051	3.554	1.791	-0.071	0.036
Sugar	SB	0.298	0.072	3.613	1.971	-0.072	0.028
Silver	SI	0.448	0.056	2.324	2.214	-0.021	0.072
Wheat	W	0.243	0.044	4.718	1.943	-0.087	0.037

Table A3: Percentage of times each commodity appear in the High4 and Low4 portfolios sorted by factors

This table presents the percentage at which each commodity is included in the High4 and Low4 portfolios, which are sorted based on benchmark factors (Basis, Momentum, and Basis-momentum) and Media emotion intensity. At the end of each month, commodities are sorted using their respective factor values. The High4 portfolio consists of the four commodities with the highest factor values, while the Low4 portfolio consists of the four commodities with the lowest factor values.

Asset Name	CRB ticker	Basis		Momentum		Basis-Momentum		Emotion intensity	
		Low 4	High 4	Low 4	High 4	Low 4	High 4	Low 4	High 4
Ethanol	AK	34.21%	0.75%	0.00%	20.68%	5.64%	43.61%	0.00%	9.77%
Aluminum	AL	2.63%	0.00%	2.26%	4.14%	0.00%	4.14%	0.00%	4.14%
Soybean Oil	BO	2.26%	0.00%	9.77%	6.77%	0.75%	5.26%	39.47%	9.40%
Corn	C	4.89%	52.26%	20.68%	9.40%	18.80%	1.13%	61.28%	0.00%
Cocoa	CC	10.15%	5.64%	25.19%	23.31%	9.77%	12.78%	5.26%	1.50%
Crude Oil	CL	21.43%	4.14%	15.79%	21.80%	12.78%	12.03%	0.00%	19.55%
Cotton	CT	19.55%	21.43%	22.93%	10.15%	37.59%	15.41%	10.90%	0.00%
Palm Oil	CU	4.14%	1.13%	3.76%	0.00%	0.38%	5.64%	0.00%	0.00%
Cattle	FC	24.06%	7.89%	8.65%	12.41%	26.69%	12.41%	6.39%	1.50%
Gold	GC	6.39%	0.00%	1.50%	16.92%	0.00%	16.92%	0.00%	40.60%
Copper	HG	15.04%	0.00%	4.51%	16.17%	0.38%	11.28%	1.50%	4.14%
Heating Oil	HO	17.67%	2.26%	17.29%	21.43%	5.64%	19.17%	10.15%	7.89%
Coffee	KC	3.76%	46.62%	31.95%	10.15%	7.14%	2.26%	1.13%	6.77%
Hogs	LH	29.32%	51.88%	28.95%	7.89%	71.80%	13.53%	78.95%	0.00%
Natural Gas	NG	16.54%	27.82%	39.47%	13.16%	47.37%	11.65%	0.00%	10.90%
Orange Juice	OJ	14.29%	23.31%	24.81%	22.18%	6.02%	25.19%	32.71%	10.53%
Palladium	PA	12.03%	0.38%	15.79%	38.72%	1.13%	21.05%	0.75%	66.92%
Platinum	PL	19.92%	0.00%	1.13%	10.53%	0.00%	14.66%	0.00%	66.54%
Brent Crude	QA	22.18%	1.50%	16.17%	22.56%	3.01%	16.92%	0.38%	52.63%
Gasoline	RB	41.35%	10.53%	5.26%	29.32%	16.92%	37.59%	1.88%	1.50%
Rough Rice	RR	7.14%	42.11%	25.56%	6.39%	42.48%	9.77%	9.40%	7.52%
Canola	RS	10.90%	11.28%	16.92%	18.42%	19.17%	14.66%	53.38%	0.38%
Soybeans	S	19.55%	0.00%	3.01%	15.79%	7.52%	19.17%	60.53%	0.00%
Sugar	SB	32.33%	31.58%	23.31%	18.80%	38.35%	32.71%	2.63%	0.00%
Silver	SI	5.64%	0.00%	7.14%	16.92%	0.00%	9.77%	0.00%	77.82%
Wheat	W	2.63%	57.52%	28.20%	6.02%	20.68%	11.28%	23.31%	0.00%

Table A4: **Percentage of overlapping between High4 and Low4 portfolios sorted by factors**

This table displays the average percentage of overlap between two portfolios: HIGH4 and LOW4. These portfolios are sorted based on media and benchmark factors. At the end of each month, commodities are sorted using factor values, and the four with the highest and the four with the lowest factor values are allocated into either the HIGH4 or LOW4 portfolios, respectively. The monthly overlapping percentage between each pair of portfolios is calculated by dividing the number of common commodities by four. The average overlapping percentage for the entire sample is determined by calculating the average of the monthly overlapping percentages.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Emotion intensity	High4	(1)	100.0									
	Low4	(2)	0.0	100.0								
Media coverage	High4	(3)	11.9	4.3	100.0							
	Low4	(4)	27.7	23.8	0.0	100.0						
News sentiment	High4	(5)	34.1	9.6	8.0	100.0						
	Low4	(6)	25.6	15.6	15.8	13.7	100.0					
Basis	High4	(7)	3.8	28.2	13.7	7.7	12.3	100.0				
	Low4	(8)	12.5	16.4	17.3	12.6	17.1	13.3	0.0	100.0		
Momentum	High4	(9)	18.0	12.7	16.9	17.3	25.6	11.7	4.2	33.0	100.0	
	Low4	(10)	10.4	20.5	17.2	15.9	8.8	21.0	42.0	6.1	0.0	100.0
Basis-momentum	High4	(11)	16.0	16.4	14.7	17.5	17.4	13.6	6.2	39.0	33.6	100.0
	Low4	(12)	2.5	24.4	18.5	3.6	7.0	17.2	27.3	18.7	22.7	0.0

Table A5: Media coverage and the cross-section of commodity portfolio returns

This table reports the cross-section tests for media coverage and other benchmark factors for the sample from February 1998 to February 2020 at the portfolio level. Model (1) presents the test for the factors introduced by Bakshi et al. (2019) and Boons and Prado (2019). Models (2) and (3) test whether the risk of media emotion intensity nearby and spreading returns are priced in the cross-section of portfolio returns against the factors in Model (1). Model (4) presents the test for the factors introduced by Szymanowska et al. (2014) and Boons and Prado (2019). Models (5) and (6) test whether the risk of media emotion intensity nearby and spreading returns are priced in the cross-section of portfolio returns against the factors in Model (4). This analysis used 42 portfolios: including (1) 30 portfolios by sorting each factor of basis, momentum, basis-momentum, media emotion intensity, and media coverage into three groups (High4, Mid, Low4) and separating to nearby and spreading strategies; (2) 12 portfolios by the two above strategies for commodity types (Energy, Grains, Livestock, Metals, Softs and Food oils). Then, the study conducts the cross-section regression of the average returns of portfolios against their exposures to factors. The estimated risk price of each factor is reported with the Newey West t-statistic from Fama-Macbeth estimation in the parentheses underneath and the t-statistic after correcting standard error using Shanken (1992)) method in the square brackets.

MODEL	Portfolio return					
	Boons and Prado (2019) and Bakshi et al. (2017)		Boons and Prado (2019) and Szymanowska et al. (2014)			
	1	2	3	4	5	6
Basis (nb)	-0.011** (-2.402) [-1.757]	-0.011** (-2.485) [-1.808]	-0.012** (-2.581) [-1.836]	-0.007* (-1.727) [-1.258]	-0.008* (-1.891) [-1.354]	-0.009** (-2.224) [-1.515]
Basis (High4 - sp)	0.000 (0.092) [0.080]	0.000 (0.103) [0.089]	0.000 (0.348) [0.292]			
Basis (Low4 - sp)	-0.004** (-2.490) [-2.039]	-0.004** (-2.478) [-2.050]	-0.004** (-2.428) [-1.987]			

Average	0.003	0.003	0.003	0.003
	(1.145)	(1.160)	(1.141)	(1.141)
	[0.845]	[0.855]	[0.806]	[0.806]
Momentum (nb)	0.011**	0.011**	0.010**	0.010**
	(2.294)	(2.300)	(2.085)	(2.085)
	[1.557]	[1.565]	[1.362]	[1.362]
Basis-momentum (nb)	0.012**	0.012**	0.010**	0.009**
	(2.355)	(2.414)	(2.634)	(2.378)
	[1.745]	[1.755]	[1.865]	[1.748]
Basis-momentum (sp)	0.001	0.001	0.000	0.001
	(0.673)	(0.666)	(0.559)	(0.559)
	[0.595]	[0.587]	[0.473]	[0.473]
Media coverage (nb)	-0.002	-0.002	-0.003	-0.001
	(-0.535)	(-0.470)	(-0.767)	(-0.411)
	[-0.379]	[-0.329]	[-0.573]	[-0.290]
Media coverage (sp)	-0.002	-0.002	-0.003**	-0.003**
	(-1.022)	(-1.022)	(-2.535)	(-2.535)
	[-0.848]	[-0.848]	[-2.088]	[-2.088]
Constant	-0.001	-0.001	-0.001***	-0.001***
	(-1.228)	(-1.436)	(-5.218)	(-4.470)
	[-1.051]	[-1.243]	[-5.110]	[-4.136]
Observation	11,130	11,130	11,130	11,130
R-Squared	0.430	0.483	0.517	0.602

Table A6: Media coverage and the cross-section of individual commodity returns

This table reports the cross-section tests for media coverage and other benchmark factors for the sample from February 1998 to February 2020 at the individual commodity level. Model (1) presents the test for the factors introduced by Bakshi et al. (2019) and Boons and Prado (2019). Models (2) and (3) test whether the risk of media emotion intensity nearby and spreading returns are priced in the cross-section of portfolio returns against the factors in Model (1). Model (4) presents the test for the factors introduced by Szymanowska et al. (2014) and Boons and Prado (2019). Models (5) and (6) test whether the risk of media emotion intensity nearby and spreading returns are priced in the cross-section of portfolio returns against the factors in Model (4). To run the cross-section tests at the commodity level, this analysis employs Fama-Macbeth (1973) regression for the nearby and spreading returns of 26 commodity futures. In the first stage, this study runs the rolling time-series regression with a one-year rolling of daily returns. In the second stage, this study runs the monthly cross-section regression of nearby and spreading returns of commodity futures on their time-varying exposures to factors. The table reports the estimated risk prices of factors with their t-statistic corrected using the Newey West method for one lag in the parenthesis and their t-statistic corrected using Shanken, 1992 method in the square brackets.

MODEL	Individual commodity futures returns					
	Boons and Prado (2019) and Bakshi et al. (2017)			Boons and Prado (2019) and Szymanowska et al. (2014)		
	1	2	3	4	5	6
Basis (nb)	-0.007 (-1.140) [-0.976]	-0.010 (-1.442) [-1.216]	-0.010 (-1.478) [-1.170]	-0.006 (-0.928) [-0.747]	-0.006 (-0.969) [-0.772]	-0.007 (-1.077) [-0.813]
Basis (High4 - sp)				0.002 (1.407) [1.236]	0.002 (1.226) [1.060]	0.001 (1.073) [0.866]
Basis (Low4 - sp)				0.001 (0.470) [0.405]	0.001 (0.652) [0.557]	0.002 (0.956) [0.746]

Average	0.004 (1.419)	0.003 (1.325)	0.003 (1.244)	
Momentum (nb)	[1.010]	[0.938]	[0.839]	
	0.002	-0.000	0.000	
	(0.353)	(-0.001)	(0.055)	
	[0.273]	[-0.001]	[0.041]	
Basis-momentum (nb)	0.015***	0.014**	0.014**	0.011**
	(2.770)	(2.563)	(2.530)	(2.083)
	[2.179]	[1.998]	[1.871]	[1.623]
Basis-momentum (sp)	0.003	0.003	0.004**	0.003
	(1.500)	(1.598)	(2.083)	(1.565)
	[1.348]	[1.427]	[1.713]	[1.240]
Media coverage (nb)		-0.009**	-0.010***	-0.004
		(-2.394)	(-2.870)	(-1.300)
		[-1.797]	[-2.050]	[-0.891]
Media coverage (sp)			-0.003**	-0.003*
			(-2.535)	(-1.933)
			[-2.146]	[-1.670]
Constant	-0.001***	-0.001***	-0.001***	-0.001
	(-3.467)	(-3.343)	(-2.795)	(-1.644)
	[-3.386]	[-3.286]	[-2.532]	[-1.452]
Observation	12,083	12,083	12,083	12,083
R-squared	0.412	0.464	0.503	0.417
			0.368	0.459

Table A7.1: Media emotion intensity versus hedging pressure

This table shows how hedging pressure interacts with media emotion intensity in explaining returns in the commodity futures market. We measure hedging pressure as the number of short positions minus the number of long positions of commercial traders scaled by total interest. Following Kang et al. (2020), we smooth hedging pressure using a one-year moving average of monthly hedging pressure. Panel A presents the correlation of hedging pressure with media emotion intensity and other benchmarks. The correlation is calculated for two settings: time-series and cross-section. We follow Boons and Prado (2019) to calculate time-series correlation using the median of time-series correlations of individual commodity hedging pressure with the other factor. The cross-section correlation is the median of the monthly cross-section correlations between hedging pressure and media emotion intensity of commodities. Panel B reports the result of the spanning test for media emotion intensity and the basis-momentum factor adding hedging pressure to the interacting factor list. We first test for only hedging pressure in the first model, only other factors in the second model and combine hedging pressure and other factors in the third model. The standard errors of estimated coefficients are all corrected using the Newey West method with one lag when doing the spanning test. The t-statistic of alpha is reported in the parenthesis. Panel C reports the cross-section test that adds hedging pressure to the factor list. We employ two-stage Fama-Macbeth (1973) regression for 42 factor-sorted and commodity-type portfolio returns. We correct the standard errors using Shanken (1992) and report the corresponding t-statistic in the parenthesis. Coefficients with "****", "***", "**", and "*" are statistically significant at 1%, 5% and 10% levels, respectively. The results are estimated for the sample from February 1998 to February 2020.

Panel A: Correlation between hedging pressure and other factors						
	Correlation	Basis	Momentum	Basis-Momentum	Emotion	Buzz
Hedging pressure	Time-series	-0.17	0.47	-0.13	0.03	0.05
	Cross-section	-0.20	0.30	0.15	0.37	-0.11
Panel B: Spanning test for media emotion intensity and basis-momentum using hedging pressure						
	Media emotion			Basis-Momentum		
Model	(1)	(2)	(3)	(4)	(5)	(6)
Alpha	0.0763*	0.1178**	0.0774*	0.0798*	0.1004**	0.0785*
t-stat (Alpha)	(1.7863)	(2.5620)	(1.8254)	(1.7719)	(2.2523)	(1.8018)
Hedging pressure (nb)	0.5348***		0.5104***	0.3251***		0.2719***

Emotion (nb)					0.0080	-0.0658
Basis-Momentum (sp)						
Basis (nb)		0.0078	-0.0565			
Momentum (nb)		-0.1234*	-0.0573		-0.1747**	-0.1467*
Average (nb)		-0.0973	-0.0008		-0.1962	-0.0411
		0.3978***	0.2013		0.4667***	0.3178**
R-squared	0.2351	0.0692	0.2548	0.0858	0.0990	0.1431
GRS-F	3.3679	6.8975	3.5016	3.0404	4.8760	3.0991
p-value	0.0665	0.0086	0.0613	0.0812	0.0272	0.0783
Panel C: Cross-section test for hedging pressure risk						
Model	(1)	(2)	(3)	(4)	(5)	(6)
Average (nb)	0.004 (1.062)	0.003 (0.865)	0.004 (1.083)	0.003 (0.853)	0.004 (1.051)	0.004 (1.052)
Hedging pressure (nb)	0.017** (2.580)		0.012* (1.769)		0.016** (2.381)	0.007 (0.938)
Emotion (nb)		0.013** (2.232)	0.013** (2.109)			0.012** (2.079)
Basis-momentum (nb)				0.013** (2.132)	0.008 (1.394)	0.010* (1.710)
Basis (nb)						-0.005 (-0.852)
Momentum (nb)						0.007 (1.031)
cons	-0.001*** (-4.756)	-0.001*** (-5.170)	-0.001*** (-4.761)	-0.001*** (-5.401)	-0.001*** (-4.686)	-0.001*** (-4.589)
R-squared	0.393	0.413	0.492	0.347	0.446	0.659

Table A7.2: Market volatility and average volatility in predicting media emotion intensity portfolio returns

This table reports the results of the time-series regression for the predictive role of market volatility and average volatility on returns of portfolios sorted by media emotion intensity. We follow Boons and Prado (2019) to measure market volatility as the sum of squared daily returns of the market portfolio (equal-weighted portfolio of commodity futures) and average volatility as the mean of monthly volatility of daily returns of individual commodity futures. Panel A shows the predictive power of market volatility for four portfolios constructed based on media emotion intensity: High4, Low4, Mid, and High4-Low4. Panel B reports the results of similar tests for average volatility. For each test, we employ the method seen in Boons and Prado (2019) that regresses the portfolio returns on a one-month lag of the volatility using WLS regression weighted on the inverted standard deviation of monthly returns calculated for the recent 12 months. The results are reported for the sample from February 1998 to February 2020.

VARIABLES	Nearby return (t + 1)					
	Emotion (High4)	Emotion (Mid)	Emotion (Low4)	Emotion (nb)		
Market volatility (t)	-2.8236 (-0.8857)	-3.4915 (-1.1022)	-2.5856 (-1.1995)	-0.2380 (-0.1130)		
R ²	0.0064	0.0191	0.0074	0.0000		
Average volatility (t)	-0.9480 (-0.6040)	-1.5765 (-1.0590)	-1.1989 (-1.1072)	0.2510 (0.2186)		
R ²	0.0024	0.0127	0.0053	0.0001		

Table A7.3: Cross-section tests for market volatility and average volatility.

This table reports the results of the cross-section tests for market volatility and average volatility. We employ the two-stage Fama-Macbeth (1973) regression to estimate the premium of each factor. In the first stage, we run the time series regression for the whole sample to estimate the time-series exposures of each portfolio to the factors. In the second stage, we run the cross-section regression of portfolio returns on these exposures. The standard errors of all estimated coefficients are adjusted following Shanken (1992) correction. We report the estimated premiums and their Shanken t-statistic in the parenthesis underneath each estimated coefficient for each model. For each selected factor, we run (1) the test for average (market) factor and the factor, (2) model adding media emotion intensity factor to the first model, and (3) model adding basis-momentum, basis and momentum factor to the second model. We follow Boons and Prado (2019) to measure market volatility as the sum of squared daily returns of the market portfolio (equal-weighted on commodity futures) and average volatility as the average monthly volatility of daily returns of commodity futures. Coefficients with "***", "**", and "*" are statistically significant at 1%, 5% and 10% levels, respectively. The results are estimated for the sample from February 1998 to February 2020.

Model	(1)	(2)	(3)	(4)	(5)	(6)
Average (nb)	0.003 (0.959)	0.003 (0.956)	0.003 (0.875)	0.003 (0.866)	0.003 (0.792)	0.003 (0.794)
Market volatility	-0.072 (-1.418)		-0.019 (-0.451)		-0.004 (-0.094)	
Average volatility		-0.028 (-0.342)		-0.012 (-0.142)		-0.011 (-0.124)
Emotion (nb)			0.013** (2.224)	0.013** (2.240)	0.012** (2.090)	0.012** (2.082)
Basis-momentum (nb)					0.010* (1.823)	0.011* (1.830)
Basis (nb)					-0.005 (-0.923)	-0.005 (-0.910)
Momentum (nb)					0.008 (1.171)	0.008 (1.176)
cons	-0.002*** (-5.662)	-0.002*** (-6.034)	-0.001*** (-5.329)	-0.001*** (-5.189)	-0.001*** (-4.737)	-0.001*** (-4.702)
R-squared	0.332	0.324	0.458	0.461	0.653	0.652

Table A7.4: Cross-section tests for funding liquidity and market liquidity risk.

This table reports the results of the cross-section tests for market liquidity and funding liquidity risk. We employ the two-stage Fama Macbeth regression to estimate the risk price of each factor. In the first stage, we run the time series regression for the whole sample to estimate the time-series exposures of each portfolio to the factors. In the second stage, we run the cross-section regression of portfolio returns on these exposures. The standard errors of all estimated coefficients are adjusted following Shanken (1992) correction. We report the estimated risk prices and their Shanken t-statistic in the parenthesis underneath each estimated coefficient for each model. For each selected factor, we run (1) the test for average (market) factor and the factor, (2) model adding media emotion intensity factor to the first model, and (3) model adding basis-momentum, basis and momentum factor to the second model. The table shows the result of testing funding illiquidity and market illiquidity. We employ two proxies for funding liquidity: Ted spread (the gap between 3-month LIBOR rate and 3-month Treasury rate) and VIX (CBOE volatility index). Coefficients with "***", "**", and "*" are statistically significant at 1%, 5% and 10% levels, respectively. The results are estimated for the sample from February 1998 to February 2020.

Model	(1)	(2)	(3)	(4)	(5)	(6)
Average (nb)	0.003 (0.738)	0.003 (0.897)	0.003 (0.711)	0.003 (0.897)	0.003 (0.716)	0.003 (0.698)
Ted spread	0.227* (1.656)		0.181 (1.365)		0.119 (0.845)	
Δ VIX		-2.736* (-1.834)		-0.927 (-0.772)		1.770 (1.228)
Emotion (nb)			0.013** (2.048)	0.012** (2.096)	0.012** (2.065)	0.013** (2.088)
Basis-momentum (nb)					0.010* (1.670)	0.011* (1.776)
Basis (nb)					-0.006 (-0.941)	-0.004 (-0.702)
Momentum (nb)					0.008 (1.111)	0.008 (1.057)
cons	-0.001*** (-4.789)	-0.002*** (-4.613)	-0.001*** (-4.353)	-0.001*** (-4.894)	-0.001*** (-4.539)	-0.001*** (-4.100)
R^2	0.344	0.360	0.480	0.465	0.659	0.640

Table A7.5: Cross-section tests for equity market risk.

This table reports the results of the cross-section tests for equity market risk and downside equity market risk. We employ the two-stage Fama-Macbeth (1973) regression to estimate the risk price of each factor. In the first stage, we run the time series regression for the whole sample to estimate the time-series exposures of each portfolio to the factors. In the second stage, we run the cross-section regression of portfolio returns on these exposures. The standard errors of all estimated coefficients are adjusted following Shanken (1992) correction. We report the estimated risk prices and their Shanken t-statistic in the parenthesis underneath each estimated coefficient for each model. For each selected factor, we run (1) the test for average (market) factor and the factor, (2) model adding media emotion intensity factor to the first model, and (3) model adding basis-momentum, basis and momentum factor to the second model. We examine the equity market risk using value-weighted CRSP excess return and also consider the downside equity market risk as suggested by Boons and Prado (2019). Coefficients with "****", "****", and "****" are statistically significant at 1%, 5% and 10% levels, respectively. The results are estimated for the sample from February 1998 to February 2020.

Model	(1)	(2)	(3)	(4)	(5)	(6)
Average (nb)	0.003 (0.891)	0.003 (0.791)	0.003 (0.855)	0.003 (0.824)	0.003 (0.777)	0.003 (0.734)
CRSP excess return	0.031 (1.519)	0.037 (1.604)	0.020 (1.065)	0.023 (1.197)	-0.005 (-0.285)	-0.012 (-0.757)
CRSP excess return (downside)		0.027 (1.620)		0.015 (1.142)		-0.013 (-1.286)
Emotion (nb)			0.012* (1.926)	0.012* (1.859)	0.012** (2.084)	0.012** (1.971)
Basis-momentum (nb)					0.011* (1.846)	0.010* (1.695)
Basis (nb)					-0.005 (-0.876)	-0.006 (-0.951)
Momentum (nb)					0.008 (1.172)	0.008 (1.096)
cons	-0.002*** (-4.841)	-0.002*** (-4.041)	-0.001*** (-4.712)	-0.001*** (-4.334)	-0.001*** (-4.676)	-0.001*** (-4.148)
R ²	0.331	0.395	0.468	0.520	0.630	0.666

Table A7.6: Principal component analysis of macroeconomic variables

This table summarizes the results of the Principle component analysis of macroeconomic variables. We employ 11 macroeconomic variables, including CPI, CPI for commodities, 3-month LIBOR, real M1, real M2, PPI for final demand of commodities, PPI for the US, PPI for total manufacturing, Unemployment, US Dollar index and Michigan consumer sentiment. All macroeconomic variables are used in difference to ensure their stationary. We conduct the principal component analysis (CPA) to extract orthogonal factors from the set of variables with varimax rotation. Panel A reports the result of PCA results with eigenvalues of adding factors. Panel B shows the loadings of factors on variables.

Panel A: Factor analysis and Eigenvalues

Factor	Eigenvalue	Difference	Proportion	Cumulative
Factor1	5.1804	3.9672	0.5180	0.5180
Factor2	1.2132	0.1263	0.1213	0.6394
Factor3	1.0869	0.2425	0.1087	0.7480
Factor4	0.8444	0.0266	0.0844	0.8325
Factor5	0.8178	0.4094	0.0818	0.9143
Factor6	0.4084	0.1809	0.0408	0.9551
Factor7	0.2276	0.1077	0.0228	0.9779
Factor8	0.1198	0.0418	0.0120	0.9898

Panel B: Factor loadings and Uniqueness

Variable	Factor1	Factor2	Factor3	Uniqueness
Δ Consumer sentiment		0.7059		0.2015
Δ CPI (commodity)	0.9269			0.1360
Δ CPI	0.9323			0.1266
Δ M1 (real)			0.6593	0.2817
Δ M2 (real)	-0.7818			0.2453
Δ PPI (commodity - final demand)	0.9278			0.1175
Δ PPI (commodity - US)	0.9115			0.1452
Δ PPI (total manufacturing)	0.9294			0.1234
Δ Unemployment		-0.7273		0.3183
Δ US Dollar index			-0.6823	0.4575

Table A7.7: Cross-section tests for macroeconomic factors.

This table presents the cross-section tests for macroeconomic factors in the model with other benchmark factors. We employ 11 macroeconomic variables, including CPI, CPI for commodities, 3-month LIBOR, real M1, real M2, PPI for final demand of commodities, PPI for the US, PPI for total manufacturing, unemployment, US Dollar index and Michigan consumer sentiment. All macroeconomic variables are used in difference to ensure their stationary. First, we conduct Principal Component Analysis (PCA) to extract the main orthogonal factors from the macroeconomic variables. After that, we use the factors predicted from PCA analysis to test their risks to the cross-section of commodity futures returns. In the first four models, we include only the average (market) factor and each macro factor. In the following four models, from Model (5) to Model (8), we add media emotion intensity factors to the set of factors employed in Model (1) to Model (4), respectively. In the last four models, Model (9) to Model (12), we further add other benchmark factors to the models of the macro factor and media factors models. We use the two-stage Fama Macbeth regression with Shanken (1992) correction to run the cross-section test. The t-statistic of each estimated risk price is calculated based on the Shanken (1992) standard error and reported in the parenthesis.

Model	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Average (nb)	0.003 (0.907)	0.003 (0.763)	0.003 (0.924)	0.003 (0.865)	0.003 (0.865)	0.003 (0.870)	0.003 (0.797)	0.003 (0.769)	0.003 (0.791)
Macro factor 1	0.253 (1.113)			-0.199 (-0.786)			-0.041 (-0.159)		
Macro factor 2		-0.902* (-1.797)			0.083 (0.196)			-0.295 (-0.634)	
Macro factor 3			-0.159 (-0.569)			0.082 (0.309)			-0.018 (-0.068)
Emotion (nb)				0.014** (2.364)	0.013** (2.223)	0.013** (2.294)	0.012** (2.151)	0.012** (2.047)	0.012** (2.085)
Basis-Momentum (nb)							0.010* (1.773)	0.010* (1.718)	0.011* (1.813)
Basis (nb)							-0.006 (-0.979)	-0.005 (-0.853)	-0.005 (-0.892)
Momentum (nb)							0.008 (1.150)	0.008 (1.139)	0.008 (1.185)
cons	-0.002*** (-5.532)	-0.002*** (-4.529)	-0.002*** (-6.002)	-0.001*** (-4.962)	-0.001*** (-4.991)	-0.001*** (-5.327)	-0.001*** (-4.510)	-0.001*** (-4.434)	-0.001*** (-4.925)
R2	0.398	0.349	0.351	0.499	0.463	0.478	0.661	0.647	0.647