Journal of Futures Markets

WILEY

RESEARCH ARTICLE OPEN ACCESS

News Sentiment and Commodity Futures Investing

Chi Yeguang[1] | Lina El-Jahel[1] | Thanh Vu[2]

[1]The University of Auckland, Auckland, New Zealand | [2]Auckland University of Technology, Auckland, New Zealand

Correspondence: Lina El-Jahel (l.eljahel@auckland.ac.nz)

Received: 23 December 2024 | **Revised:** 23 June 2025 | **Accepted:** 8 July 2025

Funding: The authors received no specific funding for this work.

Keywords: commodity futures pricing | factor investing | long-short strategy | media news sentiment

ABSTRACT

We investigate the role of media news sentiment in commodity futures investing. The weekly rebalanced long-short portfolio sorted by news sentiment generates a significant average annualized return of around 8.3% after transaction costs. The time-series spanning test reveals that the abnormal return of the long-short portfolio sorted by news sentiment is statistically significant at above 7% even after controlling for various benchmark factors. The premium of the news sentiment factor is also significantly priced at above 8% in the cross-section of commodity futures returns. Furthermore, we show that news sentiment enhances the performance of commodity futures investment portfolios.

JEL Classification: G11, G13, G14 and G17

1 | Introduction

Commodity futures returns have become an important part of investors' portfolios since financialization (Adams and Glück 2015; Belousova and Dorfleitner 2012; Daskalaki and Skiadopoulos 2011). Therefore, it is essential for investors to understand the price movement of commodity futures and their risks. This understanding helps investors form effective trading strategies. The existing literature on factors affecting the cross-section of commodity futures returns has identified some notable predictors. For example, (Miffre and Rallis 2007) shows that the momentum factor can explain the premium in commodity futures markets. (Szymanowska et al. 2014) document that the cross-section of commodity futures returns can be explained by a single basis factor. (Bakshi et al. 2017) propose a factor model with three major factors: the average factor from (Yang 2013), the basis factor from (Szymanowska et al. 2014), and the momentum factor from (Miffre and Rallis 2007). (Boons and Prado 2019) propose another salient factor, basis-momentum, which adds more predictive power to average, basis, and momentum factors. Besides the benchmarks factor suggested in (Boons and Prado 2019), we also consider other factors suggested in the literature, such as skewness (Fernandez-Perez et al. 2018), volatility (Gorton et al. 2013; Szymanowska et al. 2014), value (Asness et al. 2013), open interest (Hong and Yogo 2012; Szymanowska et al. 2014), net trade of commercial and noncommercial traders (Kang et al. 2020), as well as short-term hedging pressure and moving average hedging pressure (Basu and Miffre 2013).

(Shiller 2003) argues that the movement of the market is influenced by news. When investors read the content of a news item, the words used in this news item can impact their judgments, even when they recognize some exaggerations from the writers. As media news is an important source of information for investors, the rapid development of media news has encouraged researchers to explore how media news can impact the movement of asset risk and return. (Tetlock 2007) and (García 2013) have both examined the impact of news sentiment on market indices and found that news sentiment has a significant effect on market returns on a daily basis. However, their research only focuses on the behavioral biases induced by news sentiment. In this regard, any shift in the news sentiment can trigger a quick reaction from investors, affecting asset prices in the short term.

The literature has suggested the importance of sentiment in the commodity market. In terms of time-series impact, (Smales 2014) and (Brandt and Gao 2019) have provided evidence that the returns of gold futures and crude oil are influenced by the commodity-specific news sentiment for these commodities. In terms of cross-sectional impact, (Fernandez-Perez et al. 2020) have demonstrated that the fear of hazards poses a risk to the cross-section of commodity futures returns. (El-Jahel et al. 2024) has investigated and shown the impact of novel and old news sentiment on commodity futures returns at daily frequency. Additionally, (Fan et al. 2023) have investigated commodity-specific Twitter sentiment and found that a long-short portfolio sorted by this social media sentiment generates high and significant average returns.

In this study, we test the predictability of news sentiment for the cross-section of commodity futures returns at a weekly frequency using the long-short strategy. Moreover, we assess the portfolio performance via time-series spanning tests against the performance of long-short portfolios sorted by other factors. If the long-short portfolio sorted by news sentiment generates a significant average return that is not explained by other benchmark factors, it suggests that investors could potentially earn a premium from news sentiment investing in the commodity futures market. We investigate it by using the Fama–Macbeth regression to test the premium of the news sentiment factor.

We use the Refinitive Marketpsych Indices (RMI) data to measure weekly news sentiment for 24 commodities. The advantage of this data set is its extensive coverage of financial media outlets, providing media news sentiment data for a variety of commodities. The data set provides sentiment analysis at the level of individual news articles. It allows for aggregating sentiment across all news items each week to calculate weekly news sentiment.[1] We use the week from the end of Friday of the previous week to the end of Friday of this week to calculate the weekly news sentiment and the value of other factors at weekly frequency. To ensure that the long-short strategy is applicable, we construct the portfolios at the end of the following Monday (one trading day gap) and hold the positions for a week.

First, we investigate the returns of a long-short portfolio sorted by media news sentiment. Every week, from Friday to Friday, we divide commodities into four quartile portfolios based on the value of weekly aggregated news sentiment. Portfolios Q1, Q2, Q3, and Q4 contain the commodities in the first, second, third, and fourth quartiles of weekly news sentiment, respectively. We then create a Q4-Q1 portfolio by longing the Q4 portfolio and shorting the Q1 portfolio. Each portfolio is held for 1 week. We also consider transaction costs following the approach in (Paschke et al. 2020). Our findings reveal that the long-short portfolio sorted by news sentiment generates an average annualized return of 8.280% (t-stat = 2.173) after transaction costs. The performance of the news sentiment portfolio is comparable to those sorted by basis (10.387%), momentum (10.561%), and basis-momentum (9.175%). Similarly, the Sharpe and Sortino ratios of the news sentiment portfolio (0.450 and 0.730, respectively) are also comparable with those of basis (0.532 and 0.749), momentum (0.476 and 0.685), and basis-momentum (0.535 and 0.833).

Second, we test whether the long-short strategy based on news sentiment can be explained by other factors. We conduct a time-series spanning test, wherein we regress the return of the long-short portfolio sorted by news sentiment on the return of portfolios sorted by other factors, including basis, momentum, basis-momentum, skewness, value, volatility, change of open interest, commercial and noncommercial net trade, weekly hedging pressure and moving average hedging pressure. Although we find significant correlations between the returns of the long-short portfolio sorted by news sentiment and those sorted by momentum and moving average hedging pressure, the R-squared values of all models in the spanning test do not exceed 12%, including the model with all examined factors. Moreover, the alpha calculated from the spanning test is statistically significant across all models, generally above 7.5%. This implies that news sentiment can predict the cross-sectional return of commodity futures, and its impact is largely independent of other factors.

Third, we examine whether the premium of news sentiment is priced in the cross-section of commodities futures returns. To do so, we employ the Fama–Macbeth regression at the portfolio level. Each factor in the cross-section test is measured as the return of the Q4-Q1 portfolio sorted by that factor. In the model with only news sentiment, the estimated premium of news sentiment is statistically significant, annualized at 12.24%. Even when accounting for various factor combinations, the premium of news sentiment remains significantly priced in the commodity portfolio returns. In the model with all factors, the premium of news sentiment is 8.174% annualized. The findings suggest that investors are compensated for investing in portfolios sorted by news sentiment.

Finally, we examine whether news sentiment can improve the performance of a long-short strategy when using other factors. We combine news sentiment with other factors through two approaches. First, commodities are double-sorted by news sentiment and each factor independently. We start by dividing commodities into two portfolios by news sentiment, based on the median value of news sentiment each week. Then, we divide the commodities into two portfolios based on each factor and the median of these factors each week. The long–short portfolio sorted by each factor mostly performs better when considering only high-news-sentiment commodities than when considering all commodities. Further, when longing commodities with high-news-sentiment and high-factor values and shorting those with low-news-sentiment and low-factor values (High-High - Low-Low portfolio), this portfolio performs significantly better and has a higher risk-adjusted return (Sharpe ratio) than the long-short portfolio sorted by each factor alone. The most notable portfolios are those sorted by news sentiment and each of the basis, momentum, basis-momentum, and skewness factors. These portfolios generate Sharpe ratios of 0.624, 0.478, 0.779, and 0.739 and Sortino ratios of 0.936, 0.751, 1.184, and 1.136, respectively. In contrast, the Sharpe and Sortino ratios are lower for the portfolios single-sorted by basis (0.532 and 0.749), momentum (0.476 and 0.685), basis-momentum (0.535 and 0.833), and skewness (0.714 and 1.192).

In the second approach, commodities are ranked weekly based on news sentiment and individual factors simultaneously.

Factors are assigned in a way that a higher rank indicates higher expected returns. An equal-weighted combined rank is formulated by integrating news sentiment and other factors. Utilizing this combined rank, commodities are sorted into quartiles, and the top six commodities form the Q4 portfolio, while the bottom six form the Q1 portfolio. After that, we track the performance of the portfolio that longs the Q4 portfolio and shorts the Q1 portfolio. The findings show that portfolios sorted by combined ranking signals generate significantly higher risk-adjusted returns compared to those sorted by individual factors, except when combined with momentum. When compared to the long-short portfolio sorted by each factor, the combined ranking portfolio increases the Sharpe ratio from 0.532 to 0.566 for basis, 0.535 to 0.763 for basis-momentum, 0.714 to 0.822 for skewness, and 0.491 to 0.575 for moving average hedging pressure. This suggests that news sentiment can enhance the performance of a long-short strategy when combined with other factors.

We contribute to the existing literature in several ways. First, we investigate and suggest the predictability of news sentiment for the cross-section of commodity futures returns at a weekly frequency. (Fernandez-Perez et al. 2023) adjust the traditional signals up and down based on news sentiment values to capture the attraction of these signals, finding that sentiment-enhanced signals can improve traditional signal performance. However, our study focuses on the out-of-sample predictability of media news sentiment. We show that news sentiment can predict the cross-section of weekly commodity futures returns and generate a significant average return of the long–short strategy sorted by news sentiment.

Second, our research also adds to the literature on factor investing in the commodity futures market (Daskalaki et al. 2014; Gorton and Rouwenhorst 2006; Kang et al. 2020; Sakkas and Tessaromatis 2020). We test and show that the premium of news sentiment is significant in both the time-series spanning test and the Fama–Macbeth cross-section test. This study extends the literature of understanding the factors that their premium are priced in the cross-section of commodity futures returns (Asness et al. 2013; Bakshi et al. 2017; Boons and Prado 2019; Fernandez-Perez et al. 2018; Gorton and Rouwenhorst 2006; Gorton et al. 2013; Hong and Yogo 2012; Kang et al. 2020; Szymanowska et al. 2014). We test media news sentiment premium when controlling for these benchmark factors suggested by the literature. Our results show that the premium of media news sentiment remains significant and can not be subsumed by other benchmark factors in the time-series spanning test and the Fama–Macbeth cross-section test.

Third, our results contribute to the literature on the value of media factors in investment. We find that news sentiment can enhance the efficiency of the long–short strategy in commodity futures investing. (Fernandez-Perez et al. 2023) use RavenPack News Analytics to measure news sentiment and use this factor to overlay the key benchmark characteristics of commodities. They conclude that news sentiment significantly improves the performance of the long-short strategy by adjusting the signal value based on media optimism or pessimism towards commodities. Our research utilizes a different media data set and examines the combination of news sentiment with fundamental factors. We show that combining news sentiment with other signals also leads to better investment outcomes. Further, our study aligns with (Kang et al. 2020) in examining the impact of factors at a weekly frequency and adds to the literature on the impact of media news on asset returns (Engelberg and Parsons 2011; García 2013; Smales 2014; Tetlock 2007; Ahern and Sosyura 2015).

The rest of the paper is organized as follows. Section 2 describes the data sources and variables. Section 3 tests the value of the news sentiment signal through the long-short strategy and the time-series spanning test. Section 4 tests the risk price of news sentiment in the cross-section of commodity futures returns. Section 5 examines whether news sentiment can enhance the performance of the long-short strategy. Section 6 concludes.

2 | Data and Methodology

This section describes the data sources used in this study. The two primary sources are commodity futures (Barchart. com) and media sentiment data (Refinitiv Marketpsych indices). In addition to the data description, this section also provides definitions and methods for measuring the variables examined in subsequent analyses.

2.1 | Commodity Futures Data

This study collects data on commodity futures prices from Barchart. com. We survey 24 commodity futures that are traded on the largest futures exchanges in the United States, including CBOT, NYMEX, COMEX, and ICE. The sample period for data collection ranges from January 1998 to December 2021, aligning with the availability of media sentiment data. The specific features of the futures contracts are presented in Table 1.

We define the first nearby futures contract as the contract that is closest to maturity and still tradable within the next 2 months, in line with the definition provided by (Boons and Prado 2019). The selection of the first nearby contract is particularly relevant for the analysis of commodity futures performance at a weekly frequency, as it helps avoid the rolling of contracts when holding futures contracts for one to 4 weeks. (Szymanowska et al. 2014) note that commodity futures prices tend to exhibit wild movements when approaching the expiration date. Therefore, it is necessary to examine the movement of commodity futures returns before the last trading month to disregard the unusual trading activity during this period. Apart from the first nearby contract, the second contract is defined as the nearest-to-maturity contract following the first one, while the third and fourth nearby contracts (used in the robustness tests) are the nearest-to-maturity contracts after the second and third nearby contracts, respectively.

2.1.1 | Returns of Commodity Futures

To compute the weekly returns of futures contracts, we adopt the approach in (Boons and Prado 2019) but modify it to a

TABLE 1 | Commodity futures contracts' specification.

TRMI id	Asset name	Type	CRB ticker	Name of contract	Contract size	Exchange
BIOETH	Ethanol	Energy	AK	Ethanol (pit)	29,000 U.S. gallons	CBOT
SOIL	Soybean Oil	Food_oils	BO	Soybean Oil (pit)	60,000 pounds	CBOT
COR	Corn	Grains	C	Corn (pit)	5000 bushels	CBOT
COC	Cocoa	Softs	CC	Cocoa	10 metric tonnes	ICE
CRU	Crude Oil	Energy	CL	Crude Oil West Texas Intermediate	1000 U.S. barrels (42,000 gallons)	NYMEX
COT	Cotton	Softs	CT	Cotton #2	50,000 pounds (~ 100 bales)	ICE
CTTL	Cattle	Livestocks	FC	Feeder Cattle	50,000 pounds	CME
GOL	Gold	Metals	GC	Gold 100-oz	100 fine troy ounces	COMEX
CPPR	Copper	Metals	HG	High Grade Copper	25,000 pounds	COMEX
HOIL	Heating Oil	Energy	HO	New York Harbor ULSD	42,000 gallons	NYMEX
COF	Coffee	Softs	KC	Coffee C Arabica	37,500 pounds (~ 250 bags)	ICE
HOGS	Hogs	Livestocks	LH	Lean Hogs	40,000 pounds	CME
NGS	Natural Gas	Energy	NG	Natural Gas	10,000 MMBtu	NYMEX
ORJ	Orange Juice	Softs	OJ	Orange Juice [FCOJ-A]	15,000 pounds	ICE
PALL	Palladium	Metals	PA	Palladium	100 troy ounces	NYMEX
PLAT	Platinum	Metals	PL	Platinum	50 troy ounces	NYMEX
NSEA	Brent Crude	Energy	QA	Brent Crude Oil	1000 barrels (42,000 gallons)	NYMEX
MOG	Gasoline	Energy	RB	Gasoline Blendstock New York Harbor	42,000 gallons	NYMEX
RICE1	Rough Rice	Grains	RR	Rough Rice	2000 cwt	CBOT
RAPOIL	Canola	Food_oils	RS	Rapeseed Canola	20 tonnes	ICE
SOY1	Soybeans	Grains	S	Soybean (pit)	5000 bushels	CBOT
SUG	Sugar	Softs	SB	Sugar #11	112,000 pounds (50 long tonnes)	ICE
SLVR	Silver	Metals	SI	Silver 5,000-oz	5000 troy ounces	COMEX
WHT	Wheat	Grains	W	Chicago Soft Red Winter Wheat	5000 bushels	CBOT

Note: This table presents information regarding the 24 futures contracts used in this study, with the most active corresponding futures contract for each commodity being selected. The TRMI id, which stands for the identification code of the commodity in the media news sentiment data (Refinitiv Marketpsych Indices), and Asset name, which represents the name of the commodity used in the analyses, are included in this table. The selected commodities were classified into six distinct categories, namely energy, metal, grains, softs, food oils, and livestock. The CRB ticker, which denotes the commodity code in the Barchart. com data (formerly known as Commodity Research Bureau, CRB), is also provided. These CRB tickers are the first letters of the corresponding futures contract symbols listed on the exchanges. The last three columns contain information about the name of the selected contracts, the corresponding contract size, and the exchange where the contracts are listed.

weekly frequency. Specifically, let $F_{i,t}^{T_n}$ represent the futures price of the n^{th} nearby contract of commodity i in week t. The weekly excess return of a fully collateralized futures position is calculated as follows:

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

Our default weekly period spans from the daily settlement time on the previous week's Friday to the daily settlement time on the current week's Friday. If a commodity futures is not traded on Friday, we use the nearest tradable day before that Friday to collect the end-of-week values for each factor. To ensure the feasibility of the trading strategy, we wait one trading day after observing the weekly factor value (from Friday to Friday) and analyzing the returns for the following week (from Monday to Monday).

This study focuses on analyzing first nearby contracts due to their typically higher liquidity than further future maturing contracts. Table 2 presents the descriptive statistics of returns for each commodity futures' first nearby contracts. Generally, these contracts' average returns are close to zero with a higher

TABLE 2 | Descriptive statistics of commodity futures returns and news sentiment.

	News sentiment				Commodity futures return			
Ticker	Mean	Std.	Min	Max	Mean	Std.	Min	Max
AK	−0.018	0.148	−1.000	1.000	0.282	3.711	−18.033	16.442
BO	−0.065	0.230	−1.000	1.000	0.012	3.179	−13.227	13.452
C	−0.064	0.050	−0.500	0.124	−0.063	3.728	−16.39	20.776
CC	−0.038	0.075	−0.667	0.175	0.049	4.069	−20.833	20.243
CL	−0.118	0.051	−0.299	0.056	0.161	4.659	−27.872	23.534
CT	−0.058	0.051	−0.259	0.155	−0.038	3.519	−13.883	18.414
FC	−0.126	0.051	−0.304	0.088	0.03	2.308	−12.714	12.628
GC	−0.035	0.065	−0.361	0.141	0.132	2.412	−9.644	13.239
HG	−0.077	0.079	−0.300	0.277	0.18	3.521	−22.62	14.986
HO	−0.153	0.146	−0.643	0.410	0.138	4.294	−21.387	15.969
KC	−0.033	0.063	−0.236	0.155	−0.152	4.406	−14.307	29.95
LH	−0.123	0.075	−0.346	0.184	−0.092	3.677	−24.786	17.284
NG	−0.079	0.059	−0.294	0.132	−0.223	5.654	−21.105	23.579
OJ	−0.096	0.132	−0.629	0.667	0.009	4.218	−16.605	18.095
PA	−0.077	0.204	−0.800	0.612	0.338	4.924	−37.563	42.184
PL	−0.005	0.136	−0.550	0.451	0.159	3.277	−17.132	18.885
QA	−0.158	0.155	−0.628	0.392	0.099	4.354	−22.761	21.042
RB	−0.111	0.060	−0.296	0.109	0.264	4.654	−31.802	24.757
RR	−0.059	0.052	−0.242	0.097	−0.125	3.15	−13.379	13.324
RS	−0.084	0.088	−0.400	0.333	0.006	2.497	−13.649	9.222
S	−0.068	0.054	−0.333	0.126	0.121	3.187	−14.57	14.976
SB	−0.071	0.046	−0.339	0.370	0.045	4.154	−17.475	14.315
SI	−0.018	0.097	−0.406	0.306	0.158	4.16	−27.38	17.352
W	−0.085	0.048	−0.264	0.087	−0.131	3.902	−17.121	16.457

Note: This table presents the descriptive statistics of the weekly returns of the first nearby futures contracts corresponding to each commodity in the sample and their corresponding news sentiment. For each commodity, the table reports the mean, standard deviation, minimum, and maximum returns of the first nearby futures contract. It also provides the mean, standard deviation, minimum, and maximum of the weekly aggregated news sentiment for each commodity. The week period used in this table is defined as the period from the daily settlement time on Friday of the previous week to the daily settlement time on Friday of the week. For Fridays when the market does not open, the period from the settlement time of the closest trading day before the Friday of the previous week to the settlement time of the closest trading day before the Friday of the week is used.

standard deviation. The minimum and maximum returns for each commodity are similar in magnitude and greatly exceed the standard deviations, implying high volatility and risk in commodity futures returns. In the context of commodity futures investments, returns may be significantly derived from fluctuations in prices. Therefore, it is essential to carefully consider the selection of suitable commodities for short-term investments.

2.1.2 | Benchmark Factors

The literature on the commodity futures market primarily focuses on defining and testing risk factors on a monthly frequency. The most prominent risk factors for the commodity futures market on a monthly frequency are basis (Szymanowska et al. 2014), momentum (Bakshi et al. 2017), and basis-momentum (Boons and Prado 2019). In addition to these benchmark factors, we also consider other factors that have been well-investigated in the literature. Specifically, we measure skewness (Fernandez-Perez et al. 2018), volatility (Gorton et al. 2013; Szymanowska et al. 2014), value (Asness et al. 2013), open interest (Hong and Yogo 2012; Szymanowska et al. 2014), net trade of commercial and noncommercial traders (Kang et al. 2020), as well as short-term hedging pressure and moving average hedging pressure (Kang et al. 2020). We measure skewness using the same approach in the literature as the skewness of return for the current year. For volatility and value, modifications are made to accommodate a weekly frequency. Volatility is calculated as the standard deviation of weekly returns for the current year, while the value factor is measured as the average weekly return for the period ranging from 5.5 years to 4.5 years ago.

We measure open interest, net trade of commercial and noncommercial traders, short-term hedging pressure, and moving-average hedging pressure by collecting data from the Commitments of Traders (COT) reports released by the Commodity Futures Trading Commission (CFTC). The COT reports are generally released on Fridays, detailing the trader positions at the end of Tuesday of the same week (or the closest trading day before Tuesday in the case that the market closes on Tuesday). To ensure that the data for these factors is available at the end of the week for forming investment signals, we measure these factors for a specific week using the latest information released before the end of the week. Following the literature (Hong and Yogo 2012; Szymanowska et al. 2014), we measure the change in open interest from the previous week to the current week. For the net trade of commercial and noncommercial traders, we measure these factors following the approach of (Kang et al. 2020). Kang et al. (2020) argue that the short-term hedging pressure of commercial traders reflects commercial traders' liquidity provision to noncommercial traders rather than their hedging motives, and that the long-term (moving average) hedging pressure is a more accurate measure of the true hedging pressure in the market. Hence, we calculate weekly hedging pressure or short-term liquidity provision of commercial traders to noncommercial traders in accordance with (Kang et al. 2020).

2.2 | Media News Data and News Sentiment Measure

This study uses the Refinitiv Marketpsych Indices (RMI) to acquire the sentiment of media news towards each commodity. The data fields in RMI data are calculated from a media news storage system called Refinitiv machine-readable news (MRN). MRN has the advantage of covering a vast number of newswires, which provides an extensive range of common financial outlets. Additionally, news items fed into the system have a delay of no more than 500 ms. Using news items stored in MRN, Refinitiv calculates the sentiment of the news related to each asset by lexical analysis with a frequently updated and tested dictionary for sentiment. The study aggregates minute-level sentiment from the RMI data to weekly news sentiment. Minute-level sentiment measures the average sentiment of all news items related to a specific asset released within a minute period. The introduction of the RMI data and the linguistic analysis techniques used to measure sentiment are discussed in this study's appendix.

The MRN system collects news items related to a specific asset over a given period. Each word directly related to the asset and its associated attributes is then stored as a variable in the system. The system scores all variables obtained from the scanned text. The details on how Thomson Reuters defines and scores variables are provided in Section A1 of the Appendix. Finally, the system sums up the scores of all variables associated with the asset and stores this value in an indicator called "*buzz*". To demonstrate how "*buzz*" is calculated, we begin with "*buzz*" for a 1-min period. For a certain asset a and a specific 1-min period, the system records all variables (words and phrases with their attributes), which are denoted by the symbol $V(a)$. The system then scores each variable 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 1 min is calculated as the sum of $S(v)$ for all variables in $V(a)$.

$$buzz_t^{1-\min}(a) = \sum_{v \in V(a)} S(v) \tag{2}$$

The news sentiment data in the Refinitiv Marketpsych Indices (RMI) is reported in the Coordinated Universal Time (UTC) zone. To ensure consistency, all timestamps in the TRMI data are converted to the Central Standard Time (CST) zone. For each week T, the *buzz* of an asset a is calculated as the sum of all 1-min *buzz* for the asset a. This calculation encompasses the period from the daily settlement time of the last trading day of the previous week $T-1$ to the daily settlement time of the last trading day of the current week T. It is important to note that the more news content that is related to a specific asset, the more variables are recorded in the system. Consequently, "*buzz*" will take on a higher value.

2.2.1 | Media News Sentiment

The term "news sentiment" is frequently used to refer to two sentimental aspects of media news, positive and negative. (García 2013) and (Tetlock 2007) 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

$$I_{sen}(t,v) = \begin{cases} +1 if & vconveysapositiveviewontheasset \\ -1 if & vconveysanegativeviewontheasset \end{cases} \tag{3}$$

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

$$Newssentiment \quad (a) = \frac{\sum_{v \in V(a)} (I_{sen}(v) \times S(v))}{buzz(a)}. \tag{4}$$

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 the 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. We aggregate the weekly news sentiment from minute-level sentiments.[2] This measure considers the coverage of the related news content in each news item.

Further, this study uses long-short strategies to evaluate the value of factors. TRMI data only provides the net news sentiment.[3] We argue that net sentiment better reflects the media's overall view of an asset. Therefore, we use the net news sentiment to rank the commodity futures and test the performance of the long–short portfolio sorted by this factor.

Table 2 presents the aggregated weekly news sentiment values for the Friday-to-Friday period across all commodities.[4] The negative mean values observed for all commodities suggest that commodities tend to have a negative tone of media news on average. However, the mean values are marginally different from zero, implying that the difference between negative and positive tones of the media news is not substantial.

3 | Long–Short Portfolio Sorted by News Sentiment

This section examines the predictive power of news sentiment for commodity futures returns by constructing long-short portfolios based on this media factor. The analysis involves sorting by news sentiment and benchmark factors, followed by a performance evaluation of these portfolios. Furthermore, a time-series spanning test is conducted to explore whether news sentiment can produce excess returns that are not explained by other factors.

3.1 | Univariate Sorting and Performance of Long–Short Portfolios

The study begins by examining 24 commodities through univariate sorting using news sentiment and other benchmark factors, including basis, momentum, basis-momentum, skewness, value, volatility, change of open interest, net trade of commercial traders (commercial net trade), net trade of noncommercial traders (Noncommercial net trade), weekly hedging pressure and 52-week moving average hedging pressure (MA hedging pressure). At the end of each week, the aggregated news sentiment is calculated for all news items released from the last day's settlement time of the previous week to the last day's settlement time of the current week. For each week, commodities are sorted into four portfolios based on the quartile levels of news sentiment. Specifically, the portfolios Q1, Q2, Q3, and Q4 include the commodities in the first, second, third, and fourth quartiles of news sentiment each week. The same sorting approach is conducted for the other benchmark factors. Following this, a long–short portfolio (Q4-Q1) is formed for each factor by longing the Q4 portfolio and shorting the Q1 portfolio.

We sort commodities at the end of each week, but skip one trading day before starting to trace portfolio performance. In this case, portfolio performance is traced from the end of the next trading day to 1 week later. More specifically, factors are sorted on Fridays, and the corresponding portfolios will be constructed from the end of the next Monday to the end of the Monday a week after that. This study inverts the values of basis, skewness, volatility, change of open interest, and net trade of noncommercial traders and uses these inverted factors instead for sorting. This allows the portfolio with a higher factor value to have a higher average return than those with a lower value of the factor across all factors. Table 3 presents the returns of long-short portfolios sorted by news sentiment.

The Q4-Q1 portfolio, sorted by news sentiment, generates a high and statistically significant average annualized return of 10.246% (t-stat = 2.690). Among the portfolios sorted by news sentiment, only the Q4 portfolio has a significant average return of 8.443% (t-stat = 2.509). In comparison, Table 3 also reports the performance of long-short portfolios sorted by basis, momentum, basis-momentum, skewness, and moving average hedging pressure.[5]

Among the examined factors, the inverted basis and inverted skewness generate significant average returns. This implies that commodities with the highest values of these factors tend to generate lower returns compared to those with the lowest values of these factors. The long–short strategy has the highest performance with skewness, basis, and momentum. Skewness is the factors that bring investors the highest average returns, with an average annualized return of 14.636% (t-stat = 3.605). Momentum and basis generate comparable average returns compared to news sentiment, with average annualized returns of 11.165 (t-stat = 2.427) and 11.387 (t-stat = 2.570), respectively. Basis-momentum and moving average hedging pressure have slightly lower average returns than news sentiment. Both are around 9% annualized.

We also consider the transaction cost in the performance of these long-short portfolios. Following (Paschke et al. 2020), we measure the transaction cost of the i th nearby contract of commodity m at week $t+1$ as follows:

$$TC_{t+1}^{(m,i)} = \frac{10 + Tick^{(m)} \times CM^{(m)}}{F_{t+1}^{(m,i)} \times CM^{(m)}} \tag{5}$$

Where $Tick^{(m)}$ presents the minimum tick size of commodity m on the relevant exchange, and $CM^{(m)}$ is the contract multiplier of commodity m. We deduct the transaction costs from the portfolio returns and report the result in Table 3. The results show that, after transaction costs, the average return of the news sentiment portfolio decreases to 8.280% annualized but remains significant with a t-statistic of 2.173. The long–short portfolios sorted by basis, momentum, basis-momentum, skewness, and moving average hedging pressure also decline slightly to 10.387% (t-stat = 2.570), 10.561% (t-stat = 2.297), 9.175% (t-stat = 2.585), 14.002% (t-stat = 3.450), and 9.436% (t-stat = 2.317), respectively.[6]

In terms of the Sharpe ratio, the risk-adjusted return of news sentiment in our sample stands at 0.450. This is comparable to the levels seen in basis, momentum, basis momentum, and moving average hedging pressure. Skewness outperforms all other factors with a risk-adjusted return of 0.714. Considering downside risk, skewness also has the highest Sortino ratio (1.192). Basis-momentum, basis, and news sentiment follow with the high compensation for downside risk, with Sortino ratios of 0.833, 0.749, and 0.730, respectively.

3.2 | Time-Series Spanning Test

This section uses time-series spanning tests to examine whether news sentiment can yield additional returns independent of other influences. We regress the Q4-Q1 portfolio, sorted by news sentiment, against other Q4-Q1 portfolios sorted by benchmark factors. Each portfolio has a holding period of 1 week, with returns calculated using the first nearby contract. The benchmark factors used in the test are the same as those in the single-sorting results.

TABLE 3 | The performance of portfolios sorted by news sentiment and other factors.

		Portfolios					
		Q1	Q2	Q3	Q4	Q4-Q1 before TC	Q4-Q1 after TC
News sentiment	Ave. ret.	−1.803	1.726	3.366	8.443	10.246	8.280
	t-stat	−0.463	0.513	1.048	2.509	2.690	2.173
	Sharpe	−0.096	0.106	0.217	0.519	0.557	0.450
	Sortino	−0.125	0.159	0.327	0.776	0.904	0.730
Basis	Ave. ret.	−2.070	0.341	4.576	9.014	11.084	10.387
	t-stat	−0.579	0.101	1.339	2.434	2.743	2.570
	Sharpe	−0.120	0.021	0.277	0.504	0.568	0.532
	Sortino	−0.197	0.031	0.382	0.675	0.800	0.749
Momentum	Ave. ret.	−2.060	2.147	2.072	9.105	11.165	10.561
	t-stat	−0.525	0.671	0.649	2.305	2.427	2.297
	Sharpe	−0.109	0.139	0.134	0.477	0.503	0.476
	Sortino	−0.168	0.209	0.182	0.699	0.725	0.685
Basis-momentum	Ave. ret.	−2.565	−1.940	7.892	7.224	9.789	9.175
	t-stat	−0.790	−0.554	2.443	1.973	2.758	2.585
	Sharpe	−0.164	−0.115	0.506	0.409	0.571	0.535
	Sortino	−0.260	−0.169	0.743	0.572	0.889	0.833
Skewness	Ave. ret.	−4.275	−0.831	6.194	10.361	14.636	14.002
	t-stat	−1.155	−0.252	1.923	2.753	3.605	3.450
	Sharpe	−0.239	−0.052	0.398	0.570	0.746	0.714
	Sortino	−0.342	−0.078	0.584	0.820	1.247	1.192
Moving average HP	Ave. ret.	−1.521	1.195	5.839	8.229	9.751	9.436
	t-stat	−0.478	0.325	1.583	2.070	2.395	2.317
	Sharpe	−0.101	0.069	0.335	0.438	0.507	0.491
	Sortino	−0.155	0.103	0.517	0.576	0.737	0.713

Note: This table presents the performance of a long-short portfolio, sorted by news sentiment and other factors. Each week, commodities are sorted into four portfolios based on quartile levels of each factor. Portfolios Q1, Q2, Q3, and Q4 correspond to commodities in the first, second, third, and fourth quartiles of each factor, respectively. A Q4-Q1 portfolio is constructed by taking a long position in the Q4 portfolio and a short position in the Q1 portfolio. Each portfolio formation is delayed by a day and held for a week. The last column reports the average return of the Q4-Q1 portfolios after transaction costs, while the other column shows the average returns of portfolios before transaction costs. The weekly period in this table starts from the settlement time on the preceding week's Friday (or the nearest trading day before Friday) and ends at the settlement time on the current week's Friday (or the nearest trading day before Friday).

In the initial set of tests, the return of the long-short portfolio sorted by news sentiment is regressed on the return of the long-short portfolio sorted by each factor in Table 3. Subsequently, the spanning test regresses the return of the long-short portfolio sorted by news sentiment on different sets of factors: (1) the factors with significant average long-short portfolio returns in Table 3, and (2) all factors examined in this study. All models in this spanning test have their standard errors corrected using the Newey-West method.

Table 4 presents the results of the time series spanning test. The estimated alpha values for Models 1–5 are all significant at approximately 10%, with a t-stat greater than 2 when tested against the returns of long-short portfolios sorted by each factor. Only momentum and moving average hedging pressure show significant unconditional correlations with long-short portfolio returns sorted by news sentiment. In Model 2, the alpha drops further to 8.297%, lower than the other models with a single test factor. This suggests that news sentiment might have a similar cross-sectional impact on returns as momentum. However, the R-squared values of the spanning test across all models from 1 to 5 are less than 3%, indicating that no single benchmark factor, including momentum and moving average hedging pressure, can strongly explain the return of the long–short portfolio sorted by news sentiment.

Models 6 and 7 provide further support for the above conclusion. Model 6 demonstrates the significance of alpha at approximately 10.275% when controlling for the return of the long-short portfolio sorted by the key benchmarks in Table 3. In Model 7, where other factors are added to the test, the alpha remains significant at approximately 7.214%. Additionally, the R-squared value of the model that examines all variables is only 11.8%, which implies that the examined factors collectively provide weak explanations for the return of the long-short portfolio sorted by news sentiment. These results suggest that

TABLE 4 | Spanning test for the performance of the Q4 - Q1 portfolio sorted by news sentiment.

	Return of Q4 - Q1 portfolio sorted by news sentiment						
VARIABLES	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**
Alpha	0.189**	0.156**	0.179**	0.190***	0.182**	0.183**	0.143*
	(2.518)	(2.181)	(2.407)	(2.655)	(2.431)	(2.516)	(1.898)
Basis	0.013					−0.156**	−0.176***
	(0.260)					(−2.534)	(−2.646)
Momentum		0.172***				0.197***	0.210***
		(5.289)				(4.430)	(4.311)
Basis-momentum			0.069			0.005	−0.023
			(1.523)			(0.107)	(−0.487)
Skewness				0.006		−0.025	−0.034
				(0.104)		(−0.445)	(−0.559)
Hedging pressure (MA)					0.173***	0.152***	0.209***
					(3.624)	(3.236)	(3.503)
Other factors	NO	NO	NO	NO	NO	NO	YES
Observations	1,213	1,213	1,213	1,213	1,160	1,160	1,030
R-squared	0.000	0.044	0.004	0.000	0.034	0.078	0.111

Note: This table presents the results of the spanning test for the return of the Q4-Q1 portfolio sorted by news sentiment. To conduct this test, we regress the returns of the Q4-Q1 portfolio sorted by news sentiment on the average factor (the equally weighted returns of all the examined commodity futures), and the returns of the High4-Low4 portfolios sorted by basis, momentum, basis-momentum, skewness, value, volatility, change of open interest, commercial net trade, noncommercial net trade, weekly hedging pressure (hedging pressure in the table), and 52-week moving average hedging pressure (hedging pressure (MA) in the table). Models (1) to (5) perform the spanning test against each benchmark portfolio sorted by basis, momentum, basis-momentum, skewness, and moving average hedging pressure. These are the portfolios with statistically significant returns in the long-short strategy. Model (6) performs the spanning test against all these benchmark portfolios, while Model (7) also includes portfolios sorted by other factors examined in our study. " ", " ", and " " denote statistical significance at the 1%, 5%, and 10% levels, respectively. The standard errors of the estimated coefficients are Newey-West corrected. The test week period is from the daily settlement time on Friday (or the closest trading day before Friday) of the previous week to the daily settlement time on Friday (or the closest trading day before Friday) of the current week.

news sentiment is a meaningful signal that can generate significant and high abnormal returns for the long-short strategy at a weekly frequency, which cannot be largely explained by other benchmark factors.

4 | News Sentiment and the Cross-Section of Commodity Futures Returns

The performance of the long-short strategy, using news sentiment, implies that this factor can predict the cross-section of commodity futures returns. The spanning test also shows a weak correlation between the returns of the Q4-Q1 portfolio sorted by news sentiment and those sorted by other benchmark factors. This indicates that investors may receive compensation for investing in portfolios with higher news sentiment exposure. This section will explore whether the premium associated with news sentiment is priced into the cross-section of the weekly portfolio return. We follow (Boons and Prado 2019) and (Fernandez-Perez et al. 2018) to conduct the cross-section test at the portfolio level. The return of the Q4-Q1 portfolio sorted by a factor is used as a measure of the risk factor in the cross-section test. In the first stage of the Fama–Macbeth regression, we conduct a time-series regression of portfolio returns on the factors for each portfolio to obtain the factor loadings. These factor loadings represent the sensitivities of the portfolios to the factors. More specifically, the first stage runs the following regression:

$$R_{i,t} = \alpha_i + \sum \beta_i^F F_t + u_i \tag{6}$$

Here, $R_{i,t}$ represents the return of portfolio i in week t. F_t denotes the value of risk factor F in week t, while β_i^F represents the factor loadings or sensitivity of portfolio i to risk factor F. In the second stage, we conduct a cross-sectional regression of portfolio returns on the factor loadings for each week to obtain the time series of weekly factor premiums.

$$R_{i,t} = \gamma_0 + \sum \gamma_t^F \hat{\beta}_i^F + e_{i,t} \tag{7}$$

where γ_t^F is the premium of factor F for week t. To estimate the premium of a factor, we calculate the mean of the weekly factor premium values throughout the sample. Specifically,

$$RP^F = \bar{\hat{\gamma}}_t^F \tag{8}$$

Here, RP^F represents the estimated premium of factor F. To determine whether the premium of factor F is priced in the cross-section of portfolio returns, we test whether the value of RP^F (or the mean of $\hat{\gamma}_t^F$) is significantly different from zero.

Table 5 presents the results of the cross-sectional test, which examines whether the premium of news sentiment and other factors is priced in the cross-section of portfolio returns. Model 1 regresses the portfolio return on the news sentiment factor

TABLE 5 | Cross-section test for the risk price of news sentiment.

	Portfolio returns				
VARIABLES	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**
News sentiment	12.236***	12.090***	8.163**	10.000**	8.174*
	(2.892)	(2.867)	(1.999)	(2.424)	(1.895)
Average		13.895	−7.749	−15.004	−5.389
		(0.900)	(−0.485)	(−0.917)	(−0.318)
Basis			13.194***	10.663**	7.574*
			(2.966)	(2.492)	(1.691)
Momentum			13.798***	12.874***	7.445
			(2.965)	(2.701)	(1.534)
Basis-momentum			9.137**	8.535**	6.529*
			(2.523)	(2.330)	(1.693)
Skewness				14.806***	9.335**
				(3.531)	(2.157)
Moving average HP				6.400	6.768
				(1.465)	(1.462)
Constant	3.724	3.852	2.517	3.836	3.235
	(1.328)	(1.371)	(0.906)	(1.345)	(1.066)
R-squared	0.045	0.079	0.209	0.298	0.482

Note: This table presents cross-sectional tests for news sentiment and other benchmark factors for a sample from January 1998 to March 2021. To perform the cross-sectional test at the portfolio level, this study uses Fama– Macbeth (1973) regression. For each of the examined factors, we sort commodities into four portfolios (Q1, Q2, Q3, and Q4) based on the factor's quartile level every week. This creates 48 portfolios from 12 factors: news sentiment, basis, momentum, basis-momentum, skewness, volatility, value, change of open interest, net trade of commercial traders, net trade of Noncommercial traders, weekly hedging pressure, and 52-week moving average hedging pressure, for the cross-section test. In the first stage, the time-series regression of portfolio returns on factors is conducted for each commodity futures to obtain the factor loadings. In the second stage, the cross-sectional regression of portfolio returns on their factor loadings is run to obtain the factor premium for each week. Model (1) uses only news sentiment. Model (2) includes the average factor in addition to what is in Model (1). Model (3) further integrates factors from boonsBasisMomentum2019. Model (4) adds skewness and moving average hedging pressure to Model (3), while Model (5) comprises all other factors under examination. The estimated risk prices of factors are reported with their Newey West corrected t-statistic in the parenthesis. "***", "**", and "*" represent statistical significance at the 1%, 5%, and 10% levels, respectively. The reported R-squared is the average R-squared derived from the weekly cross-section regression.

only, while Model 2 adds the average factor to Model 1. The average factor acts as a market factor, which is measured as the average return of all examined commodity futures each week. Model 3 tests against news sentiment and other benchmark factors in (Boons and Prado 2019), including average, basis, momentum, and basis-momentum factors. Model 4 adds skewness and moving average hedging pressure to Model 3. Model 5 regresses the portfolio returns on all examined factors.

In each model, we report the estimated premium of each factor (RP^F) and their t-stat to determine whether the premium of each factor is significantly different from zero. The t-stat reported in parentheses is the t-stat from the one-sample t-test but with the standard errors corrected using the Newey-West method. The R-squared reported for each model is the average R-squared obtained from the weekly cross-section tests in the second stage of the Fama–Macbeth regression. In each model, we report the estimated premium of each factor (RP^F) and their t-stat to determine whether the premium of each factor is significantly different from zero. The t-stat reported are those obtained from the one-sample t-test, but with the standard errors corrected using the Newey-West method. The R-squared reported for each model is the average R-squared obtained from the weekly cross-section tests in the second stage of the Fama–Macbeth regression.

Table 5 shows that the estimated premium of news sentiment is statistically significant across all test models, with a value of 12.236% annualized in the model with only news sentiment and a value of 8.174% in the model with all factors. These results demonstrate that the premium of news sentiment is significantly priced in the cross-section of portfolio returns. Model (3) reveals that among the key factors suggested by (Boons and Prado 2019), the premiums of basis, momentum, and basis-momentum are significantly priced in the cross-section of commodity portfolio returns. The significance of these factors persists in Model (4) but markedly drops in Model (5). In the model with all examined factors, the estimated premium of momentum becomes insignificant. At the same time, the Newey-West t-stat of the basis and basis-momentum factors also decreases from more than 2.5 in Model 3 to below 1.7 in Model 5.

Skewness is a factor that contributes significantly to premiums for investors. The premium of skewness is significant in both Model 4 and Model 5. However, the premium of skewness decreases from 14.806 in Model 4 to 9.335 in Model 5. This suggests that the premium of skewness is somewhat linked to the risk of other factors. On the other hand, the premium of news sentiment decreases from Model 2 to Model 3, and further to Model 5. This indicates that the premium of news sentiment may be connected to the risk represented by other factors.

Among the factors examined, news sentiment has the second highest t-stat in the cross-section test in Model 5, only surpassed by the significance of the skewness factor. In Model 1, the average R-squared is 4.5%, implying that news sentiment contributes 4.5% in predicting the cross-section of commodity portfolio returns. From Model 1 to 3, R-squared only increases from 4.5% to 20% with four more factors, including salient benchmarks such as basis, momentum, and basis-momentum. This suggests that the predictability of news sentiment is at a comparable level with these benchmark factors.

5 | News Sentiment and Portfolio Performance

Our results show that news sentiment generates a significant average return for the long-short strategy. This return is weakly explained by the return of the long–short portfolios sorted by other factors. The cross-section test also reveals that the risk of news sentiment is priced in the cross-section of portfolio returns and survives throughout the models with various risk factors. These results imply that news sentiment may have an impact on commodity futures returns that are relatively independent of other factors. Therefore, this section investigates whether using news sentiment in conjunction with other factors can enhance the performance of the long-short strategy.

5.1 | Double-Sorting on News Sentiment and Other Factors

This section aims to test the value of news sentiment in enhancing portfolio performance when sorted by other factors. To do so, we conduct a double sort based on news sentiment and each factor and examine the returns of double-sorted portfolios. The single-sorting results reveal that some factors can generate high returns for the long-short strategy. However, the high and significant returns of the Q4-Q1 portfolios in Table 3 mostly come from one side (long or short) of the portfolios. For example, high-momentum and high-basis-momentum commodities can generate significant returns, but low-momentum and low-basis-momentum do not. Given the relative independence of the news sentiment signal compared to other factors, the question now is whether news sentiment can help identify commodities with a better ability to generate high returns among those in each portfolio, sorted by other factors.

To assess the impact of news sentiment, we employ a double-sorting method, as in (Boons and Prado 2019), to categorize commodities based on news sentiment and the value of each factor. This involves individually sorting commodities by their news sentiment values and factor values weekly. With this method, commodities are divided into high and low-news-sentiment groups based on the median news sentiment value across all commodities each week. Similarly, commodities are also divided into two portfolios, high-factor and low-factor values, based on the median value of the factor for each week. After sorting commodities based on news sentiment and each factor individually, we observe the performance of different portfolios. The High & High portfolio consists of commodities with both high-news-sentiment and high-factor values. Conversely, the Low & Low portfolio includes commodities with both low-news-sentiment and low-factor values. The High-sentiment & Low-factor portfolio includes commodities with high-news-sentiment but low-factor values. Likewise, the Low-sentiment & High-factor portfolio contains commodities with low-news-sentiment but high-factor values. We also consider the High-Low portfolio, which is sorted by news sentiment. This involves longing commodities with high-news-sentiment and shorting those with low-news-sentiment within each portfolio, sorted by other factors. This helps examine the ability to differentiate commodities with varying return-generating capacities within each portfolio, sorted by other factors.

Table 6 displays the performance of a portfolio that is independently double-sorted by news sentiment and each factor. It also outlines the results of single-sorting commodities into two portfolios, High and Low, shown in the third and fourth columns. The final six columns report the returns of double-sorted portfolios. In the single-sorting results, the High–Low portfolio sorted by news sentiment yields a significant annualized average return of 5.774% (t-stat = 2.289). However, this performance is only half of the Q4-Q1 portfolio in Table 3, which is expected as the commodities in the Q4 portfolio generate much higher returns than those in the other three quartiles. Regarding other factors, only basis, momentum, basis-momentum, skewness, and moving-average hedging pressure can generate significant average returns. For all factors examined, the High-sentiment & High-factor portfolios typically generate a higher average return than the High portfolio, single-sorted by the corresponding factor.

The High–Low portfolios, sorted by news sentiment within each portfolio, single-sorted by other factors, have positive average returns, except within the portfolio of low-value factor (−2.783%). This implies that high-news-sentiment commodities can generate average returns that are no lower than those of low-news-sentiment commodities within each portfolio that is single-sorted by other factors. The significant High–Low portfolio, sorted by news sentiment, can be observed for high-basis (6.900%), low-momentum (6.923%), high-value (12.549%), low-volatility (6.744%), high-change-of-open-interest (8.800%), high-net-trade-of-commercial-traders (6.308%), high-net-trade-of-noncommercial-traders (7.242%), low-weekly-hedging-pressure (7.256%), and low-moving-average-hedging-pressure (5.846%) commodities. However, some of the above portfolios are not significant in the single sort. For example, the high-value portfolio sorted by the value factor generates an insignificant average return of 3.234%. However, within this portfolio, news sentiment can help to identify the commodities that have the ability to generate higher returns than others. This can be used to construct a long–short portfolio by news sentiment that can generate a significant return of 12.549% (t-stat = 3.249).

Additionally, the High–Low portfolios sorted by each factor within the portfolio of commodities with high-news-sentiment values mostly generate positive average returns. The only exception is the case of weekly hedging pressure, which has a return of −0.385% and a t-stat of −0.147. The results in Table 6 suggest that, in the portfolio of high-news-sentiment commodities, those with high-factor values tend to generate returns that are not lower than those with low-factor values. Similarly, in the portfolio of commodities with high-factor values, those with

TABLE 6 | Independent double-sorting on news sentiment and other factors.

		Single sorting		News sentiment					
				High		Low		High-Low	
		Ave. ret.	t-stat	Ave. ret.	t-stat	Ave. ret.	t-stat	Ave. ret.	t-stat
News sentiment	High	5.715	1.993						
	Low	−0.059	−0.019						
	High-Low	5.774	2.289						
Basis	Contango	6.726	2.141	10.003	2.902	3.103	0.822	6.900	1.974
	Backwardation	−0.911	−0.301	0.443	0.132	−3.184	−0.914	3.626	1.098
	Cont.-Back.	7.637	2.726	9.561	2.675	6.287	1.714		
Momentum	Winnners	5.394	1.732	7.609	2.200	2.564	0.696	5.045	1.506
	Losers	−0.038	−0.012	3.719	1.066	−3.204	−0.881	6.923	1.961
	Win.-Los.	5.432	1.844	3.891	1.043	5.769	1.514		
Basis-momentum	High	7.602	2.497	10.048	3.053	4.872	1.337	5.176	1.551
	Low	−2.253	−0.759	0.166	0.051	−5.157	−1.407	5.323	1.477
	High-Low	9.854	4.076	9.882	3.174	10.029	2.759		
Skewness	High	8.037	2.632	10.307	3.093	4.850	1.315	5.456	1.542
	Low	−2.687	−0.870	0.961	0.285	−4.393	−1.208	5.353	1.584
	High-Low	10.724	3.915	9.346	2.762	9.243	2.392		
Value	High	3.234	0.968	9.285	2.565	−3.263	−0.798	12.549	3.249
	Low	3.855	1.134	2.560	0.682	5.343	1.338	−2.783	−0.743
	High-Low	−0.621	−0.225	6.725	1.909	−8.606	−2.130		
Volatility	High	4.361	1.346	5.477	1.534	3.347	0.888	2.130	0.599
	Low	0.967	0.337	4.228	1.328	−2.516	−0.728	6.744	2.062
	High-Low	3.394	1.261	1.249	0.361	5.863	1.614		
Change of open interest	High	4.348	1.458	8.048	2.305	−0.752	−0.209	8.800	2.394
	Low	0.488	0.161	0.782	0.234	0.478	0.129	0.304	0.086
	High-Low	3.860	1.596	7.266	2.085	−1.229	−0.359		
Commercial net trader	High	3.202	1.070	6.124	1.814	−0.184	−0.051	6.308	1.776
	Low	1.670	0.545	3.883	1.138	−0.018	−0.005	3.901	1.079
	High-Low	1.532	0.606	2.242	0.660	−0.166	−0.046		
Noncommercial net trade	High	3.287	1.068	6.923	2.000	−0.319	−0.086	7.242	2.017
	Low	1.582	0.535	3.621	1.068	−0.089	−0.024	3.710	1.013
	High-Low	1.704	0.688	3.302	0.960	−0.230	−0.064		
Weekly hedging pressure	High	2.216	0.696	3.931	1.146	1.251	0.298	2.773	0.714
	Low	2.602	0.892	6.654	1.878	−0.603	−0.181	7.256	2.056
	High-Low	−0.385	−0.147	−2.723		−0.727	1.852	0.509	
Moving average HP	High	6.984	2.123	7.462	2.126	5.727	1.342	1.735	0.439
	Low	−0.238	−0.079	3.109	0.867	−2.718	−0.800	5.846	1.673
	High-Low	7.222	2.638	4.317	1.141	8.445	2.257		

Note: This table reports the nearby returns of portfolios that are independently double-sorted by news sentiment and other factors. First, commodities are sorted into two portfolios based on the median value of each factor at the end of each week. The return with t-statistics of the portfolio single-sorted by each factor is presented in the third and fourth columns. All the portfolios are created by sorting commodities using the factor value at the end of each week (Friday-to-Friday) and then holding for one week with a one-day lag (Monday-to-Monday). The last six columns report the returns of the portfolios sorted by news sentiment, and each reported factor independently.

high-news-sentiment values tend to generate returns that are not lower than those with low-news-sentiment values. This finding suggests that the High-High - Low-Low portfolio, which longs the commodities that have both high-factor values and high-news-sentiment values and shorts the commodities that have both low-factor values and low-news-sentiment values, may generate better returns than the High–Low portfolios single-sorted by each factor in the corresponding double sorts.

Table 7 shows the performance of the High-High - Low-Low portfolio, sorted by both news sentiment and each factor. The table shows that this portfolio generates significant average returns when sorted using news sentiment alongside basis (11.150%), momentum (8.860%), basis-momentum (13.259%), skewness (12.733%), and the moving-average hedging pressure (8.667%) after transaction costs. Furthermore, the risk-adjusted returns (Sharpe ratio) and the compensation for downside risk (Sortino ratio) of the High-High - Low-Low portfolios in Table 7 are higher than those of the Q4-Q1 portfolio sorted by each factor in Table 3. This suggests that these double-sorted portfolios not only boost the average returns of the long–short strategy but also improve the compensation for the strategy's risk.

5.2 | Combining News Sentiment and Each Factor

We test a simple method that combines news sentiment and each factor to implement a long-short strategy. Specifically, we rank commodities weekly based on their news sentiment value and factor values. For each pair of news sentiment and a factor in a specific week, the combined signal for a commodity is calculated as the sum of the ranks of that commodity according to news sentiment and the factor. This method assigns equal weights to news sentiment and each factor in the signal. To ensure the combined signal reflects the ranks of both news sentiment and the factor, we only consider commodities with data available for both elements each week.

Specifically, each week we rank commodities based on news sentiment and other factors from lowest to highest value. A lower rank score corresponds to a lower factor value. We then combine the rank of news sentiment and other factors by summing the rank scores. Afterward, we divided 24 commodities into four groups based on the quartile level of the combined rank. We identify the six commodities in the first quartile with the lowest score (Q1 portfolio) for shorting and the six commodities in the fourth quartile with the highest score (Q1 portfolio) for longing.

Table 8 reports the performance of the Q4–Q1 portfolio sorted by the combined signal of news sentiment and each factor. The left side of the table presents the performance of the portfolios, single-sorted by each factor, while the right side of the table shows the performance of portfolios sorted by the combined signal. The performance of the Q4-Q1 portfolios, single sorted by each factor in Table 8 is the average return after transaction cost of the long-short portfolio reported in Table 3. In Table 8, we report the result for the full set of test factors for comparison with the performance of the long–short portfolio sorted by combined ranks.

Using news sentiment in conjunction with other factors such as basis momentum, skewness, and moving-average hedging pressure improves returns for Q4 - Q1 portfolios. The return for

TABLE 7 | The performance of portfolios constructed by longing commodity futures with high-news-sentiment and high-factor values and shorting those with low-news-sentiment and low-factor values.

	High High - Low Low portfolios sorted by news sentiment and each factor							
	before TC				**after TC**			
Factor	**Ave. Ret**	**t-stat**	**Sharpe**	**Sortino**	**Ave. Ret**	**t-stat**	**Sharpe**	**Sortino**
Basis (inverted)	13.340	3.598	0.745	1.100	11.150	3.016	0.624	0.936
Momentum	10.927	2.883	0.597	0.950	8.860	2.309	0.478	0.751
Basis-momentum	14.347	4.117	0.852	1.304	13.259	3.760	0.779	1.184
Skewness (inverted)	15.057	4.185	0.867	1.337	12.733	3.570	0.739	1.136
Value	5.614	1.448	0.325	0.500	1.493	0.384	0.086	0.132
Volatility (inverted)	7.907	2.113	0.438	0.683	5.909	1.553	0.322	0.492
Change of open interest (inverted)	8.750	2.405	0.498	0.789	5.109	1.383	0.287	0.449
Commercial net trade	6.595	1.771	0.367	0.540	3.691	0.988	0.205	0.299
Noncommercial net trade (inverted)	7.913	2.180	0.452	0.695	4.578	1.259	0.261	0.396
Weekly hedging pressure	4.819	1.349	0.280	0.459	2.820	0.787	0.163	0.257
Moving-average HP	10.084	2.745	0.581	0.932	8.667	2.366	0.501	0.798

Note: This table reports the performance of the portfolio constructed by longing the commodity futures with high values of news sentiment and high values of the factor and shorting those with low values of news sentiment and low values of the factor. For each factor and each week, commodities are sorted into two portfolios, High and Low, based on the median value of the factor in the corresponding week. The High-High - Low-Low portfolio for a factor is formed by selecting the commodity futures appearing in both High portfolios sorted by news sentiment and the factor for longing, and selecting the commodity futures appearing the both Low portfolios sorted by news sentiment and the factor for shorting. All the portfolios are created by sorting commodities using their factor values at the end of each week (Friday-to-Friday) and then holding for one week with a one-day lag (Monday-to-Monday). We report the average returns of the long-short portfolios for both before and after transaction costs.

TABLE 8 | The performance of long–short strategy (Q4 - Q1) based on the combined rank of news sentiment and each factor.

		Single sort on each factor Q4 - Q1 after TC	Combining news sentiment and each factor Q1	Q4	Q4 - Q1 before TC	Q4 - Q1 after TC
News sentiment	Ave. ret.	8.280				
	t-stat	2.173				
	Sharpe	0.450				
	Sortino	0.730				
Basis	Ave. ret.	10.387	−2.337	9.385	11.722	10.029
	t-stat	2.570	−0.649	2.735	3.194	2.734
	Sharpe	0.532	−0.134	0.566	0.661	0.566
	Sortino	0.749	−0.201	0.820	0.993	0.849
Momentum	Ave. ret.	10.561	−1.495	7.465	8.960	7.393
	t-stat	2.297	−0.386	2.179	2.195	1.812
	Sharpe	0.476	−0.080	0.451	0.454	0.375
	Sortino	0.685	−0.115	0.670	0.695	0.572
Basis-momentum	Ave. ret.	9.175	−5.009	9.705	14.714	13.103
	t-stat	2.585	−1.372	2.975	4.141	3.687
	Sharpe	0.535	−0.284	0.616	0.857	0.763
	Sortino	0.833	−0.427	0.950	1.331	1.183
Skewness	Ave. ret.	14.002	−5.319	10.495	15.813	14.246
	t-stat	3.450	−1.430	3.176	4.405	3.970
	Sharpe	0.714	−0.296	0.658	0.912	0.822
	Sortino	1.192	−0.400	0.974	1.515	1.361

Table 8: The performance of long-short strategy (High6-Low6) based on the combined rank of news sentiment and each factor (cont.)

		Single sort on each factor Q4 - Q1 after TC	Combining news sentiment and each factor Q1	Q4	Q4 - Q1 before TC	Q4 - Q1 after TC
Value	Ave. ret.	−3.138	2.874	5.700	2.826	0.696
	t-stat	−0.837	0.712	1.584	0.744	0.183
	Sharpe	−0.188	0.160	0.356	0.167	0.041
	Sortino	−0.274	0.238	0.500	0.241	0.059
Volatility	Ave. ret.	3.737	−1.546	7.375	8.921	7.220
	t-stat	0.951	−0.441	2.121	2.414	1.954
	Sharpe	0.197	−0.091	0.439	0.500	0.405
	Sortino	0.326	−0.113	0.674	0.880	0.711
Change of open interest	Ave. ret.	4.175	1.033	7.726	6.693	4.685
	t-stat	1.183	0.283	2.355	1.925	1.347
	Sharpe	0.245	0.059	0.488	0.399	0.279
	Sortino	0.398	0.082	0.743	0.667	0.466
Commercial net trade	Ave. ret.	−0.122	0.037	6.765	6.728	4.664
	t-stat	−0.036	0.010	2.118	1.971	1.366
	Sharpe	−0.007	0.002	0.439	0.409	0.283

(Continues)

Table 8: The performance of long-short strategy (High6-Low6) based on the combined rank of news sentiment and each factor (cont.)

		Single sort on each factor	Combining news sentiment and each factor			
		Q4 - Q1 after TC	Q1	Q4	Q4 - Q1 before TC	Q4 - Q1 after TC
	Sortino	−0.012	0.003	0.652	0.640	0.442
Noncommercial net trade	Ave. ret.	−2.773	−0.050	6.426	6.477	4.437
	t-stat	−0.838	−0.014	1.992	1.928	1.321
	Sharpe	−0.174	−0.003	0.413	0.400	0.274
	Sortino	−0.255	−0.004	0.614	0.609	0.416
Weekly hedging pressure	Ave. ret.	2.693	−1.128	6.490	7.618	6.261
	t-stat	0.700	−0.345	1.923	2.151	1.767
	Sharpe	0.145	−0.072	0.398	0.446	0.366
	Sortino	0.213	−0.106	0.573	0.717	0.587
Moving average HP	Ave. ret.	9.436	−2.531	8.611	11.142	9.893
	t-stat	2.317	−0.754	2.432	3.058	2.715
	Sharpe	0.491	−0.160	0.515	0.647	0.575
	Sortino	0.713	−0.232	0.708	1.046	0.928

Note: This table presents the performance of the long-short strategy using the combined rank of each factor and news sentiment. Firstly, each commodity is ranked based on the value of news sentiment and the values of each factor. Secondly, the combined rank of news sentiment and a factor for a commodity in a specific week is defined as the sum of the ranks of the commodity on news sentiment and the factor in that week. Next, the commodities are sorted into four portfolios (Q1, Q2, Q3, and Q4) using the quartile levels of the combined rank of news sentiment and each factor. The long-short strategy is conducted by longing the commodity futures in the fourth quartile with the highest combined rank and shorting the commodities in the first quartile with the lowest combined rank for each factor. Columns 2 to 6 refer to the performance of single-sorting portfolios by each factor. The last four columns report the performance of a single-sorting portfolio by the combined rank of news sentiment and each factor.

basis-momentum increases to 13.103% (t-stat = 3.687); for skewness, it increases to 14.246% (t-stat = 3.970), and for moving-average hedging pressure, it improves to 9.893% (t-stat = 2.715) after transaction costs. However, combining news sentiment with momentum does not generate a higher return. This aligns with the spanning test result showing a significant correlation between the returns of long-short portfolios sorted by news sentiment and momentum. For factors that generate insignificant average returns in single sorts, like volatility, change of open interest, net trade of commercial traders, and net trade of noncommercial traders, their performance within a long-short strategy improves significantly when combined with news sentiment.

There is a consistent pattern when combining factors with news sentiment. Portfolios sorted by combined signals (Q4 - Q1) have significantly better Sharpe and Sortino ratios than those sorted by individual factors. This suggests that incorporating news sentiment can improve the return for each unit of risk taken, thus underlining the potential benefits of news sentiment in factor investing. Although combining news sentiment with basis does not yield a higher average return, the risk-adjusted performance metrics (Sharpe and Sortino ratios) of the portfolio, which are based on both news sentiment and basis, match the levels seen in portfolios sorted solely by basis. The data in Table 8 supports the double-sorting findings that news sentiment does add value, optimizing the long–short strategy when used alongside other factors.

6 | Conclusion

This study examines the impact of news sentiment on factor investing and commodity futures returns. First, we show that news sentiment can predict commodity futures returns. A long–short portfolio, which longs in the commodities in the highest quartile and shorts those in the lowest quartile of weekly news sentiment, generates an annualized return of around 8.280% after deducting transaction costs. This performance is at a similar level to other benchmarks in the literature, such as basis, momentum, and basis-momentum. The risk-adjusted return of the long–short portfolio sorted by news sentiment is only lower than that shorted by the skewness factor. The spanning tests also show that the return of the long-short portfolio sorted by news sentiment is at best weakly explained by other commonly considered factors.

Second, we show that, in the cross-section test, the risk of news sentiment is priced in the cross-section of commodity portfolio returns. This suggests that investors are compensated for bearing higher risk due to the higher sensitivity of their portfolios to news sentiment. The risk premium of the news sentiment factor is 12.236% in the single model. When controlling for other factors in the cross-section test, the estimated risk price of news sentiment remains statistically and economically significant at 8.174%. The results suggest that the risk price of news sentiment is not linked strongly to the risk of other factors.

Third, we also investigate whether news sentiment can enhance the performance of long-short strategies that already incorporate other factors. Our findings show that considering news sentiment can help identify commodity futures with a higher potential for returns. Commodity futures with high news sentiment tend to outperform those with low news sentiment across most factor-sorted portfolios. Building on this, we test a long–short strategy that buys commodity futures with high factor and news sentiment values and sells those with low values. This approach yields significantly higher average returns than strategies that only consider one factor. Furthermore, combining the rank of news sentiment and each factor provides a useful signal for the long–short strategy. The improved risk-adjusted returns of the combined signal-sorted portfolios indicate that news sentiment can enhance portfolio investment efficiency by enhancing the return per unit of risk.

Acknowledgments

Open access publishing facilitated by The University of Auckland, as part of the Wiley - The University of Auckland agreement via the Council of Australian University Librarians.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

Data that support the findings of this study are available from Barchart.com and Refinitiv Marketpsych Indices. Restrictions apply to the availability of these data, which were used under license for this study. Data are available from the author(s) with the permission of Barchart.com and Refinitiv Marketpsych Indices.

Endnotes

[1] The sentiment dictionary for RMI data receives periodic updates. These updates apply only to future sentiment calculations, not to historical periods. This ensures there is no look-ahead bias in RMI's news sentiment calculations.

[2] We aggregate data from the minute level to ensure we can calculate news sentiment for the exact period between the previous day's closing time and the current day's closing time. This aggregation is not affected by the absence of sentiment words in media news at the minute level, as the weekly sentiment is aggregated for the average sentiment per sentiment word, not the simple average of weekly sentiment.

[3] Thomson Reuters also provides another data set called TRNA, which includes separate positive and negative news sentiment measures. However, this data set covers a shorter period (from 2003) and a smaller number of commodities, which is not efficient for sorting the commodities and testing the performance of the long-short portfolio.

[4] In our sample, 12 commodities have sentiment words in news coverage for 100% of weeks, while 20 commodities have sentiment words for more than 99% of weeks.

[5] In our sample, the Q4-Q1 portfolios sorted by volatility, value, change of open interest, net trade of commercial traders, noncommercial traders, and weekly hedging pressure are not statistically significant at the 10% level. Therefore, we only report the factors with significant long-short portfolio performance in Table 3.

[6] We also examine the role of social media sentiment. TRMI data provides social media sentiment at the minute level, and we calculate weekly social media sentiment using the same approach as news sentiment. In our sample, the correlation between social media and news sentiment is only 0.137, suggesting that news and social media capture different aspects of sentiment. The long-short portfolio sorted by social media sentiment generates an insignificant average return of 4.573% (t-stat = 1.240) before transaction costs, and 2.407% (t-stat = 0.653) after transaction costs. Therefore, in this study, we focus only on news sentiment. All test results for social media sentiment are available upon request.

References

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

Ahern, K. R., and D. Sosyura. 2015. "Rumor Has It: Sensationalism in Financial Media." *Review of Financial Studies* 28, no. 7: 2050–2093.

Asness, C. S., T. J. Moskowitz, and L. H. Pedersen. 2013. "Value and Momentum Everywhere: Value and Momentum Everywhere." *Journal of Finance* 68, no. 3: 929–985.

Bakshi, G., X. Gao, and A. G. Rossi (2017). Understanding the Sources of Risk Underlying the Cross Section of Commodity Returns. *Management Science*. Publisher: INFORMS.

Basu, D., and J. Miffre. 2013. "Capturing the Risk Premium of Commodity Futures: The Role of Hedging Pressure." *Journal of Banking & Finance* 37, no. 7: 2652–2664.

Belousova, J., and G. Dorfleitner. 2012. "On the Diversification Benefits of Commodities From the Perspective of Euro Investors." *Journal of Banking & Finance* 36, no. 9: 2455–2472.

Boons, M., and M. P. Prado. 2019. "Basis-Momentum." *Journal of Finance* 74, no. 1: 239–279.

Brandt, M. W., and L. Gao. 2019. "Macro Fundamentals or Geopolitical Events? A Textual Analysis of News Events for Crude Oil." *Journal of Empirical Finance* 51: 64–94.

Daskalaki, C., A. Kostakis, and G. Skiadopoulos. 2014. "Are There Common Factors in Individual Commodity Futures Returns?" *Journal of Banking & Finance* 40: 346–363.

Daskalaki, C., and G. Skiadopoulos. 2011. "Should Investors Include Commodities in Their Portfolios After All? New Evidence." *Journal of Banking & Finance* 35, no. 10: 2606–2626.

El-Jahel, L., Y. Chi, and T. Vu. 2024. "Novel and Old News Sentiment in Commodity Futures Markets." *Energy Economics* 140: 108006.

Engelberg, J. E., and C. A. Parsons. 2011. "The Causal Impact of Media in Financial Markets." *Journal of Finance* 66, no. 1: 67–97.

Fan, J. H., S. Binnewies, and S. De Silva. 2023. "Wisdom of Crowds and Commodity Pricing." *Journal of Futures Markets* 43, no. 8: 1040–1068.

Fernandez-Perez, A., B. Frijns, A.-M. Fuertes, and J. Miffre. 2018. "The Skewness of Commodity Futures Returns." *Journal of Banking & Finance* 86: 143–158.

Fernandez-Perez, A., A.-M. Fuertes, M. Gonzalez-Fernandez, and J. Miffre. 2020. "Fear of Hazards in Commodity Futures Markets." *Journal of Banking & Finance* 119: 105902.

Fernandez-Perez, A., A.-M. Fuertes, J. Miffre, and N. Zhao (2023). Newswire Tone-Overlay Commodity Portfolios.

García, D. 2013. "Sentiment During Recessions." *Journal of Finance* 68, no. 3: 1267–1300.

Gorton, G., and K. G. Rouwenhorst. 2006. "Facts and Fantasies About Commodity Futures." *Financial Analysts Journal* 62, no. 2: 47–68.

Gorton, G. B., F. Hayashi, and K. G. Rouwenhorst. 2013. "The Fundamentals of Commodity Futures Returns." *Review of Finance* 17, no. 1: 35–105.

Hong, H., and M. Yogo. 2012. “What Does Futures Market Interest Tell Us About the Macroeconomy and Asset Prices?” *Journal of Financial Economics* 105, no. 3: 473–490.

Kang, W., K. G. Rouwenhorst, and K. Tang. 2020. “A Tale of Two Premiums: The Role of Hedgers and Speculators in Commodity Futures Markets.” *Journal of Finance* 75, no. 1: 377–417.

Miffre, J., and G. Rallis. 2007. “Momentum Strategies in Commodity Futures Markets.” *Journal of Banking & Finance* 31, no. 6: 1863–1886.

Paschke, R., M. Prokopczuk, and C. Wese Simen. 2020. “Curve Momentum.” *Journal of Banking & Finance* 113: 105718.

Sakkas, A., and N. Tessaromatis. 2020. “Factor Based Commodity Investing.” *Journal of Banking & Finance* 115: 105807.

Shiller, R. J. 2003. “From Efficient Markets Theory to Behavioral Finance.” *Journal of Economic Perspectives* 17, no. 1: 83–104.

Smales, L. A. 2014. “News Sentiment in the Gold Futures Market.” *Journal of Banking & Finance* 49: 275–286.

Szymanowska, M., F. De Roon, T. Nijman, and R. Van Den Goorbergh. 2014. “An Anatomy of Commodity Futures Risk Premia.” *Journal of Finance* 69, no. 1: 453–482.

Tetlock, P. C. 2007. “Giving Content to Investor Sentiment: The Role of Media in the Stock Market.” *Journal of Finance* 62, no. 3: 1139–1168.

Yang, F. 2013. “Investment Shocks and the Commodity Basis Spread.” *Journal of Financial Economics* 110, no. 1: 164–184.

Appendix

Thomson Reuters Marketpsych Indices Data and Lexical Analysis

This study uses the Refinitiv MarketPsych Indices (RMI) to calculate media news sentiment. This data source is originally called Thomson Reuters Marketpsych indices (TRMI), owned by Thomson Reuters. TRMI was developed by Refinitiv, a department within the structure of Thomson Reuters. However, Refinitiv was sold to the London Stock Exchange in 2021. Although Refinitiv is still responsible for developing TRMI data with the same approaches as conducted when under the ownership of Thomson Reuters, TRMI data is renamed to Refinitiv Marketpsy Indices to reflect the change in the data ownership.

TRMI is a system of indicators that reflect a variety of psychological aspects from news developed by the Thomson Reuters news reading system. TRMI's superiority comes from the fact that 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.

Different sources of information will affect investor sentiment in different aspects. In addition, the linguistic properties of different sources are also different. Therefore, analyzing language requires a suitable approach for each source. In the case of this study, news media and social media are two distinctly linguistic sources. If the content on news media has carefully censored content, language selection is also more complex and purposeful, language used on social media is more free and expresses emotions of the writer more directly. In social media, writers also use a variety of content expressions, such as adding emoticons or acronyms. The TRMI data uses a news reading system specifically developed to analyze the language of the press and the language of social media. Algorithms and the lexicon dictionary are also updated regularly to accommodate changes in language over time.

TRMI contains three main groups of indices. The first group reflects the psychological aspects of news, such as sentiment, optimism, fear, etc. The second group provides some metrics for macroeconomic information, and the third group covers some Buzz metrics for assets or market-moving topics related to the asset. Each index value will be calculated based on the news that appeared in a recent window and updated at a specific frequency. TRMI provides three combinations of the time window and frequency: 1 min - 1 min, 24 h - 1 h and 24 h - daily. Because each commodity futures has a specific trading time, the closing time of trading for a commodity futures might differ from that of other commodity futures. Hence, with the accuracy of measuring sentiment and emotion indices at the minute level and updating the data every minute, TRMI has an advantage of measuring precisely the aggregated sentiment and emotion indices for different time periods and frequencies.

Traditionally, textual analysis often employs the “bag of word” 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”.

TRMI calls words or phrases with their attributes “Variable”. Accounting, earning, ambiguity, and anxiety are all examples of words. In these examples, accounting and earnings 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, we will focus on an example of a sentence: “Experts expect gold will have a 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, identifying the word “higher” as an “up” word. Hence, the system will record a Variable “PriceUp f” to refer to a piece of information on a higher price 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.