

Commodity Futures Report Text Sentiment and Returns

Qian Han Pei-lin Hsieh Lu Zhang

Sun Yat-Sen University; National ChengChi University; China Construction Bank
Guangdong Branch

August 13, 2025

- Chinese commodity market and professional analysts' reports
 - We apply text analysis to a large set of analysts' **futures research reports** to construct sentiment and uncertainty indicators that quantify their tone.
- Sentiment and uncertainty indices
 - Positive and negative word counts (*POS*, *NEG*, *AGG*) and an uncertainty measure (*Uncet*).
- We show
 - Sentiment indices are strongly associated with contemporaneous returns, especially *NEG*.
 - We find no evidence that the sentiment indices predict future commodity returns.
 - Uncertainty better identifies the risk of supply–demand imbalances and helps forecast future return risk premia. [Detail in later slides]

- China's commodity market is globally important
 - 58% of global copper demand; 64% of global nickel demand; 60% of global primary aluminium production; 49% of global refined zinc production.
- **Five** commodity futures exchanges
 - Since 2016, the commodity futures and options markets have grown rapidly.
 - Over 45 commodity futures and over 38 options are listed.
 - Many contracts rank among the global top five by trading volume.
- About **150** futures brokerage firms
 - Many firms have professional teams that publish weekly research reports.
 - The Wind database aggregates most of these PDF reports.

Literature (Sentiment)

- Assessing investor sentiment at the individual-stock and market-index levels [Baker and Wurgler, 2006, Han, 2008, Connolly et al., 2005, Gao and Süß, 2015]
 - Sentiment index construction from economic indicators, Google search attention, Twitter, etc.
- Applying *textual* analysis to measure sentiment [Engelberg and Parsons, 2011, Tetlock, 2007, Tetlock et al., 2008, Engelberg et al., 2012]
 - Textual analysis of news and media to extract tone and study return predictability.

Literature (Sentiment and Returns)

- Asymmetric associations with contemporaneous returns and with earnings prediction [Huang et al., 2014, Smales, 2014, Zheng, 2015, Fernandez-Perez et al., 2020, Omura and Todorova, 2019, Maghyreh and Abdoh, 2020].
- Return predictability from sentiment is mixed across studies
 - Positive evidence: [Tetlock et al., 2008].
 - Negative evidence: [Jiang et al., 2019].
 - Conditional on setting/method: [Engelberg et al., 2012, Fan et al., 2023, Bartov et al., 2018].
- Findings vary by data source (news, social media, analyst reports), textual method, and whether past returns are controlled for.
- Our contribution:
 - New evidence based on professional analysts' research reports in Chinese commodity markets.
 - We show: **textual uncertainty is linked to a specific commodity pricing risk component.**

- Sample selection:
 - 37 actively traded futures contracts listed after March 2016 on the Zhengzhou Commodity Exchange, Dalian Commodity Exchange, and Shanghai Futures Exchange, spanning four sectors.
 - 12 agricultural, 8 chemical, 4 energy, and 13 metal futures.
- Data:
 - Analyst research reports on commodities.
 - Period: March 2016–December 2021.
- Methodology:
 - Sentiment analysis using natural language processing (NLP).
 - Sentiment measures: $POS_{i,t}$, $NEG_{i,t}$, $AGG_{i,t}$.
 - Uncertainty: $Uncert_{i,t} = |POS_{i,t}| + |NEG_{i,t}|$.

Market average AGG, POS, NEG, and Uncertainty

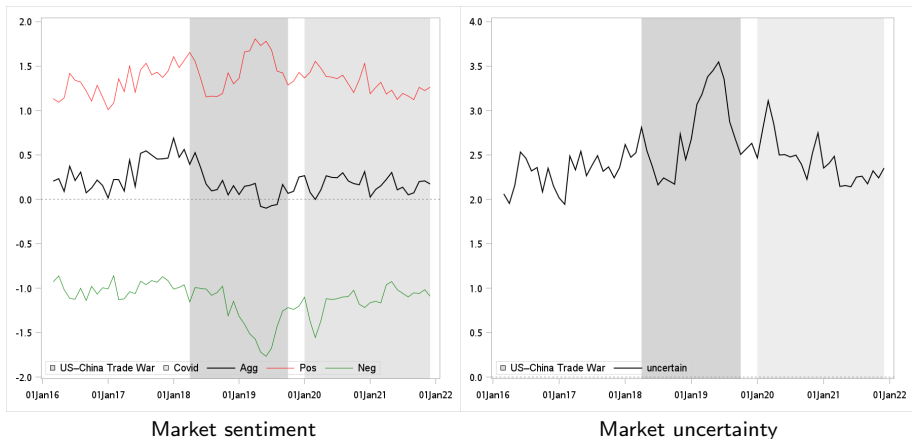


Figure: Time-series plots of Market average AGG, POS, NEG, and Uncet

AGG, POS, NEG, and Uncertainty of Soybean

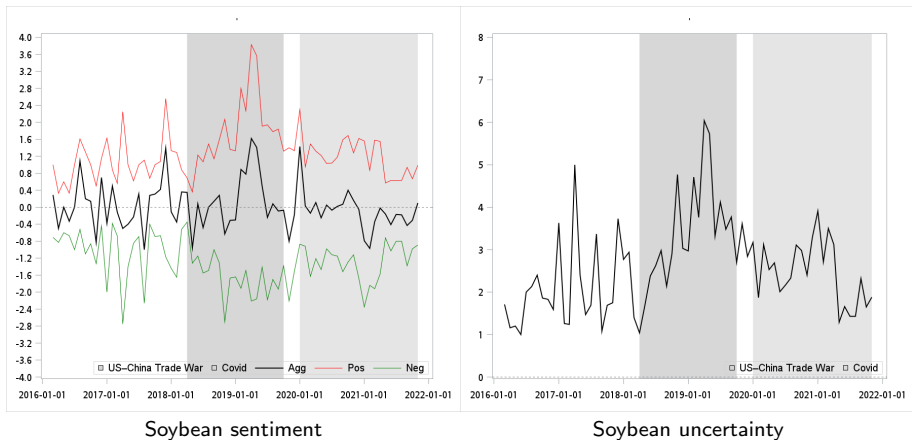


Figure: Time-series plots of Soybean AGG, POS, NEG, and Uncet

Descriptive Statistics

	AGG					Uncet			
	mean	std	min	max		mean	std	min	max
All	0.212	0.706	-1	6	All	2.485	1.184	0	10.86
Agriculture	0.05	0.516	-1	6	Agriculture	2.491	1.453	0	10.86
Chemistry	0.187	0.392	-0.84	6	Chemistry	2.374	0.873	0.33	10.86
Energy	0.27	0.32	-1	1.39	Energy	2.399	0.649	0	6
Metal	0.347	1.005	-1	6	Metal	2.569	1.218	0.16	10.86

	POS					NEG			
	mean	std	min	max		mean	std	min	max
All	1.354	0.847	0	7.17	All	-1.13	0.513	-3.69	0
Agriculture	1.284	0.898	0	7.17	Agriculture	-1.207	0.651	-3.69	0
Chemistry	1.278	0.5	0.17	7.17	Chemistry	-1.096	0.425	-3.69	-0.15
Energy	1.335	0.351	0	3.29	Energy	-1.065	0.372	-3	0
Metal	1.464	1.052	0	7.17	Metal	-1.106	0.451	-3.69	0

Text Sentiment Index V.S. Common Pricing Factors

	AGG	POS	NEG	Uncet	Mom	basismom	basisratio	OI	Liq	Vol	Skew
AGG	1.000	0.739	0.181	0.443	0.059	0.007	0.031	0.028	0.040	0.013	-0.014
P-value		0.000	0.000	0.000	0.003	0.715	0.116	0.155	0.045	0.526	0.490
POS		1.000	-0.485	0.641	-0.045	0.033	0.006	0.019	-0.008	0.006	0.025
P-value			0.000	0.000	0.022	0.097	0.771	0.327	0.673	0.753	0.201
NEG			1.000	-0.307	0.144	-0.026	0.049	0.018	0.060	0.005	-0.050
P-value				0.000	0.000	0.197	0.014	0.363	0.002	0.819	0.012
Uncet				1.000	-0.038	0.004	-0.030	0.109	0.057	0.008	-0.001
P-value					0.054	0.841	0.126	0.000	0.004	0.682	0.970
Mom					1.000	0.328	0.283	0.007	-0.029	0.029	0.158
P-value						0.000	0.000	0.712	0.144	0.137	0.000
basismom						1.000	0.577	0.031	-0.021	-0.017	0.101
P-value							0.000	0.112	0.295	0.392	0.000
basisratio							1.000	0.010	0.052	0.008	-0.081
P-value								0.630	0.009	0.693	0.000
OI								1.000	0.166	0.134	0.014
P-value									0.000	0.000	0.480
Liq									1.000	0.316	-0.028
P-value										0.000	0.158
Vol										1.000	0.008
											0.670

Text Sentiment and Commodity Futures Return

Panel Data Regression model to test linear risk premium explaining and predicting power. **senti=[AGG, POS, NEG]**

- contemporaneous returns :

$$R_{i,t} = \beta_0 + \beta_1 \text{senti}_{i,t} + \beta_2 \text{Mom}_{i,t} + \beta_3 \text{Basis}_{i,t} + \beta_4 \text{Basismom}_{i,t} + \beta_5 \Delta \text{OI}_{i,t} + \beta_6 \text{LIQ}_{i,t} + \beta_7 \text{VOL}_{i,t} + \beta_8 \text{SKEW}_{i,t} + \beta_9 \text{D_Ex}_t + \beta_{10} \text{D_M2}_t + \beta_{11} \text{Inflation}_t + \beta_{12} \text{SR}_{i,t} + \beta_{13} \text{HR}_{i,t} + \varepsilon_{i,t}$$

- next period returns :

$$R_{i,t+1} = \beta_0 + \beta_1 \text{senti}_{i,t} + \beta_2 \text{Mom}_{i,t} + \beta_3 \text{Basis}_{i,t} + \beta_4 \text{Basismom}_{i,t} + \beta_5 \Delta \text{OI}_{i,t} + \beta_6 \text{LIQ}_{i,t} + \beta_7 \text{VOL}_{i,t} + \beta_8 \text{SKEW}_{i,t} + \beta_9 \text{D_Ex}_t + \beta_{10} \text{D_M2}_t + \beta_{11} \text{Inflation}_t + \beta_{12} \text{SR}_{i,t} + \beta_{13} \text{HR}_{i,t} + \varepsilon_{i,t}$$

Text Sentiment and Commodity Futures Return (Asymmetry on contemporaneous returns)

- Aggregate sentiment and negative sentiment are sensitive to market performance, while positive sentiment is not significant.

Panel Regression

Ret_t	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$AGG_{i,t}$	0.064*** (3.08)				0.041** (2.09)			
$POS_{i,t}$		0.024 (0.97)				0.011 (0.47)		
$NEG_{i,t}$			0.074*** (3.31)				0.053** (2.52)	
$Unct_{i,t}$				0.019 (-1.09)				0.015 (0.90)
Adj R ²	0.175	0.172	0.176	0.1723	0.287	0.286	0.288	
Controls	No	No	No	No	Yes	Yes	Yes	Yes
FE&TE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Text Sentiment and Commodity Futures Return (No evidence in linear predicting power)

Ret_{t+1}	(1)	(2)	(3)	(4)
AGG	0.023 (1.21)			
POS		0.010 (0.53)		
NEG			0.017 (0.82)	
Unct				0.018 (1.17)
mom	-0.009 (-0.37)	-0.007 (-0.30)	-0.009 (-0.35)	-0.008 (-0.33)
basismom	0.018 (0.80)	0.018 (0.80)	0.018 (0.80)	0.018 (0.76)
basisratio	-0.453*** (-20.05)	-0.453*** (-20.05)	-0.453*** (-20.04)	-0.452*** (-20.00)
Adj R ²	0.344	0.343	0.343	0.344
Other Controls	Yes	Yes	Yes	Yes
FE&TE	Yes	Yes	Yes	Yes

Text Sentiment and Commodity Futures Return (Uncertainty and risk premium)

We follow Boons & Prado(2019) to focus on 3 risk.

- (1) momentum: **idiosyncratic risk** [Bakshi et al.(2019)]
- (2) basis momentum: **risk of imbalances in the supply of and demand** [Boons & Prado(2019)]
- (3) basis ratio: **hedging pressure risk**

- Regression model :

$$R_{i,t+1} = \beta_0 + \beta_1 \mathbf{Unct}_{i,t} + \beta_2 \mathbf{Mom}_{i,t} + \beta_3 \mathbf{Mom}_{i,t} \times \mathbf{Unct}_{i,t} + \beta_4 \mathbf{Basis}_{i,t} + \beta_5 \mathbf{Basis}_{i,t} \times \mathbf{Unct}_{i,t} + \beta_6 \mathbf{Basismom}_{i,t} + \beta_7 \mathbf{Basismom}_{i,t} \times \mathbf{Unct}_{i,t} + \sum [\text{Controls} + \text{Controls} \times \mathbf{Unct}_{i,t}] + \varepsilon_{i,t}$$

Text Sentiment and Commodity Futures Return (Unct better identifies the risk and improves the predicting power)

Ret_{t+1}	(1)	(2)	(3)	(4)	(5)	(6)
Unct	0 (0.03)	0.003 (0.21)	0.012 (0.77)	0.014 (0.96)	0.032* (1.73)	0.03* (1.7)
mom	0.038 (1.41)	0.016 (0.61)	0.011 (0.38)	0.04 (1.45)	0.016 (0.58)	0.008 (0.28)
mom X Unct	-0.013 (-0.89)	-0.013 (-0.82)	-0.011 (-0.75)	-0.008 (-0.49)	-0.007 (-0.4)	-0.004 (-0.23)
basismom	-0.115*** (-4.01)	-0.107*** (-3.37)	-0.069** (-2.32)	-0.124*** (-4.25)	-0.119*** (-3.71)	-0.082** (-2.71)
basismom X Unct	0.067*** (5.09)	0.074*** (5.11)	0.06*** (4.49)	0.078*** (5.6)	0.086*** (5.66)	0.071*** (5.04)
basisratio	-0.367*** (-13.53)	-0.379*** (-12.29)	-0.415*** (-14.55)	-0.377*** (-13.65)	-0.388*** (-12.46)	-0.42*** (-14.57)
basisratio X Unct	-0.031** (-2.54)	-0.035*** (-2.6)	-0.024* (-1.95)	-0.026** (-2.09)	-0.029** (-2.16)	-0.019 (-1.55)
Other Controls	N	N	N	Y	Y	Y
FE&TE	N	FE	FE&TE	N	FE	FE&TE
R ²	0.201	0.185	0.349	0.206	0.193	0.353

Text Sentiment and Commodity Futures Return (Controlling with the Trade-War and Covid-19)

- Only *basismom X Unct* and *basisratio* remain significant.
- Texture uncertainty better identifies the risk of speculators (or intermediary) who clean the imbalance position of futures contracts.

Ret _{t+1}					
Intercept	0.000 (0.00)	Unct	0.031 (1.63)		
mom	0.022 (0.480)	basismom	-0.061 (-1.410)	basisratio	-0.433*** (-10.84)
mom X Unct	-0.016 (-0.710)	basismom X Unct	0.061*** (3.450)	basisratio X Unct	-0.021 (-1.340)
mom X War	-0.054 (-0.650)	basismom X War	-0.024 (-0.320)	basisratio X War	0.050 (0.710)
mom X War X Unct	0.029 (0.620)	basismom X War X Unct	0.009 (0.250)	basisratio X War X Unct	0.008 (0.260)
mom X Covid	0.058 (0.850)	basismom X Covid	-0.148* (-1.840)	basisratio X Covid	0.001 (0.010)
mom X Covid X Unct	-0.046 (-1.06)	basismom X Covid X Unct	0.119* (1.870)	basisratio X Covid X Unct	0.017 (0.250)
Adj R ²	0.363				
Other Controls	Yes				
FE&TE	Yes				

- Textual sentiment analysis shows:
 - An asymmetric association between sentiment indices and contemporaneous returns.
 - No evidence of linear predictability for next-month returns.
- However, textual uncertainty helps identify risk:
 - Basis–momentum conditional on uncertainty is statistically significant.
 - Textual uncertainty captures risks arising from supply–demand imbalances (which is highly related to the willingness of speculators (or intermediaries) to take cross-maturity spread positions).

References

The End

- Baker, M., Wurgler, J., 2006. Investor sentiment and the cross-section of stock returns. *The Journal of Finance* 61, 1645–1680.
- Bartov, E., Faurel, L., Mohanram, P.S., 2018. Can twitter help predict firm-level earnings and stock returns? *The Accounting Review* 93, 25–57.
- Connolly, R., Stivers, C., Sun, L., 2005. Stock market uncertainty and the stock-bond return relation. *Journal of Financial and Quantitative Analysis* 40, 161–194.
- Engelberg, J.E., Parsons, C.A., 2011. The causal impact of media in financial markets. *the Journal of Finance* 66, 67–97.
- Engelberg, J.E., Reed, A.V., Ringgenberg, M.C., 2012. How are shorts informed?: Short sellers, news, and information processing. *Journal of Financial Economics* 105, 260–278.

References II

- Fan, J.H., Binnewies, S., De Silva, S., 2023. Wisdom of crowds and commodity pricing. *Journal of Futures Markets* 43, 1040–1068.
- Fernandez-Perez, A., Fuertes, A.M., Gonzalez-Fernandez, M., Miffre, J., 2020. Fear of hazards in commodity futures markets. *Journal of Banking & Finance* 119, 105902.
- Gao, L., Süß, S., 2015. Market sentiment in commodity futures returns. *Journal of Empirical Finance* 33, 84–103.
- Han, B., 2008. Investor sentiment and option prices. *The Review of Financial Studies* 21, 387–414.
- Huang, A.H., Zang, A.Y., Zheng, R., 2014. Evidence on the information content of text in analyst reports. *The Accounting Review* 89, 2151–2180.
- Jiang, F., Lee, J., Martin, X., Zhou, G., 2019. Manager sentiment and stock returns. *Journal of Financial Economics* 132, 126–149.

References III

- Maghyereh, A., Abdoh, H., 2020. The tail dependence structure between investor sentiment and commodity markets. *Resources Policy* 68, 101789.
- Omura, A., Todorova, N., 2019. The quantile dependence of commodity futures markets on news sentiment. *Journal of Futures Markets* 39, 818–837.
- Smales, L.A., 2014. News sentiment in the gold futures market. *Journal of Banking & Finance* 49, 275–286.
- Tetlock, P.C., 2007. Giving content to investor sentiment: The role of media in the stock market. *The Journal of finance* 62, 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, 1437–1467.

Zheng, Y., 2015. The linkage between aggregate investor sentiment and metal futures returns: A nonlinear approach. *The Quarterly Review of Economics and Finance* 58, 128–142.