ChatGPT and Commodity Return

Shen Gao, Ph.D.
School of Computer Science and Engineering, University of Electronic Science and Technology of China

Shijie Wang
Ph.D. candidate, School of Economics, Shandong University, China

Yuanzhi Wang
Ph.D. candidate, School of Economics, Shandong University, China

Qunzi Zhang, Ph.D.
School of Economics, Shandong University, China

Published in *Journal of Futures Markets*, March 2025, Vol. 45, No. 3 and Available at:
https://doi.org/10.1002/fut.22568

This paper investigates the ability of a ChatGPT-based indicator to forecast excess returns of a commodity futures index. Using ChatGPT to extract information from over 2.5 million articles from nine international newspapers, we demonstrate that our constructed commodity news ratio index significantly predicts future commodity returns, both in-sample and out-of-sample. Furthermore, it outperforms traditional textual analysis methods, including Bidirectional Encoder Representations from Transformers (BERT) and Bag-of-Words (BoW), while indicating economic significance within an asset allocation framework. The results highlight the critical role of ChatGPT in forecasting commodity market dynamics and provide valuable insights for both financial market participants and researchers.

Introduction

Forecasting commodity prices is important because they affect global trade, inflation, and financial markets. Commodities reflect economic conditions and often guide decisions by investors and central banks (Wang, 2001; Lutzenberger, 2014; Smales, 2014; Kilian and Zhou, 2022). Since the early 2000s, index-based investing has led to stronger links between commodity markets and other financial assets, increasing both volatility and sensitivity to global financial trends (Malik and Hammoudeh, 2007; Silvennoinen and Thorp, 2013; Mensi *et al*., 2016; Han *et al*., 2016). Many studies have explored traditional predictors like momentum (Fuertes *et al*., 2010), productivity differences (Ready *et al*., 2017), market sentiment (Omura and Todorova, 2019), and rare events (Zhang, 2021), but fewer have looked at how news content helps forecast commodity returns.

Recent research shows that text-based indicators can improve forecasts of financial outcomes (Baker *et al*., 2016; Fisher *et al*., 2022; Bybee *et al*., 2024). Large language models (LLMs) such as ChatGPT offer a new way to analyze language in news articles. While some studies have used LLMs to predict stock returns (Chen, Tang, Zhou, and Zhu, 2023; Chen, Kelly, and Xiu, 2023), few apply them to commodities. This study fills that gap by using ChatGPT to develop a Commodity News Ratio Index (CNRI) based on over 2.5 million news articles from 1946 to 2022, covering 18 major commodities from nine leading newspapers across the U.S., U.K., and Canada.

We follow a four-step method. First, we collect relevant news articles. Second, we use ChatGPT to judge whether each article suggests rising or falling commodity prices. Articles signalling increases are labelled "good news." Third, we calculate each commodity's news ratio as the difference between good and bad news, divided by the total number of articles over the past three months. Finally, we apply Partial Least Squares (PLS) to combine these ratios into one index, the CNRI, which focuses on the most useful predictive signals (Wold, 1966; Kelly and Pruitt, 2013; Huang *et al*., 2015; Zhao *et al*., 2021). Our results show that the CNRI can strongly forecast commodity futures index excess returns. In-sample, it delivers an R^2 of 2.44 on average, and the predictive power increases during the financialization period, especially when market volatility is low. The index remains effective when controlling for 14 business cycle indicators (Welch and Goyal, 2008), and it also helps forecast key economic measures like GDP, industrial production, and inflation, often outperforming alternative text-based methods like BERT or word counts (Devlin *et al*., 2018; Rogers *et al*., 2021).

This paper makes three main contributions. First, it shows how ChatGPT can be used to extract valuable insights from long-term news data to forecast commodity returns. Second, it improves text-based financial forecasting by using a large dataset and an advanced model to go beyond simple word counts. Third, it finds that commodity-related news helps explain broader economic trends, with the CNRI often signalling upcoming growth. These insights are useful for both researchers and practitioners who want to better understand the link between media content and market behavior.

Relevance of the Research Question

This research is important because it addresses a critical and underexplored area in the asset pricing literature: the predictive role of commodity-related textual information, particularly as extracted by large language models (LLMs). While existing studies have demonstrated the efficacy of financial news in forecasting equity returns (Tetlock, 2007; Manela and Moreira, 2017; Bybee *et al*., 2024; Hong *et al*., 2024), their extension to commodity markets remains limited. Given the growing financialization of commodities and their heightened sensitivity to macroeconomic and cross-asset dynamics (Henderson *et al*., 2015; Baker, 2021; Silvennoinen and Thorp, 2013), understanding how commodity news flows shape return expectations is increasingly vital. Moreover, traditional textual analysis methods, such as Bag-of-Words, fail to capture the contextual richness and semantic nuance present in high-dimensional news data. By leveraging the superior language understanding capabilities of LLMs like ChatGPT (Devlin *et al*. 2018; Rogers *et al*., 2021), this study not only enhances predictive performance but also improves the measurement accuracy of market-relevant information signals. Furthermore, while prior research has highlighted the predictive relevance of macro-financial variables (Chen, Rogoff, and Rossi, 2010; Gargano and Timmermann, 2012; Zhang, 2021), few studies have systematically combined machine learning techniques with commodity-specific news analytics. This paper fills that gap by applying a Partial Least Squares (PLS) framework (Wold, 1966; Kelly and Pruitt, 2013) to extract a robust commodity news ratio index, offering a novel pathway to forecast commodity futures returns and contribute to the broader literature on information-driven asset pricing.

Data and Methodology

This study constructs a comprehensive dataset of commodity-related news and explores its predictive power for commodity futures returns using large language models (LLMs) and a partial least squares (PLS) approach.

We begin by collecting a large corpus of news articles from nine internationally recognized newspapers, covering the period from January 1946 to December 2022. These articles are retrieved from databases such as ProQuest and Factiva. The selection of newspapers ensures a broad coverage of global commodity events and allows for the analysis of sentiment across different media sources. The news articles are filtered to focus on 18 commodities, selected based on their inclusion in the widely followed S&P GSCI index. To identify relevant articles for each commodity, we apply a classification method that combines WordNet's lexical database with manual adjustments, ensuring the coverage of both direct synonyms and semantically related terms.

For market data, we use monthly excess returns on an equal-weighted commodity futures portfolio, sourced from Levine *et al*. (2018). This portfolio includes a broad set of commodities traded on major international exchanges. The return data serve as the main target variable in our forecasting framework.

To quantify the sentiment embedded in news articles, we use two language models: ChatGPT and BERT. ChatGPT is prompted to classify news content as suggesting a rise, decline, or unclear direction in commodity prices. Based on these classifications, we construct a Commodity News Ratio (CNR) for each commodity by comparing the number of positive and negative articles over a three-month window. For BERT, we first label a sample of 5,000 articles through human annotation, categorizing them as good, bad, or neutral news. Using this labelled data, we fine-tune the BERT model to predict sentiment for the rest of the dataset.

To combine the information from multiple commodities, we adopt a two-step PLS method. In the first step, we assess the predictive strength of each individual commodity's news sentiment by linking its CNR to future commodity returns. In the second step, we aggregate these signals into a single composite index—the Commodity News Ratio Index (CNRI). This index is constructed in a way that assigns more weight to commodities whose news sentiment has historically shown stronger predictive power.

The resulting CNRI reflects the common component of commodity-related news sentiment and serves as a forward-looking indicator for excess returns in the commodity futures market. This framework allows us to systematically evaluate the role of commodity news in shaping market expectations over time, using tools from natural language processing and predictive modeling.

Empirical Results

In this section, we focus on a conventional in-sample regression for predictive analysis

$$R_{t+1:t+h} = \alpha + \beta CNRI_t + \epsilon_{t+1:t+h}, \quad (1)$$

where $R_{t+1:t+h}$ is the accumulated commodity futures index excess returns between months $t+1$ and $t+h$, $CNRI_t$ is the commodity news ratio index derived from ChatGPT, and $\epsilon_{t+1:t+h}$ is an error term.

Table 1 reports the results from in-sample predictive regressions from 1946 to 2022, where the next one-month to twelve-month accumulated commodity futures index excess returns are regressed on the commodity news ratio index (CNRI) derived from ChatGPT. The research period extends from 1946.01 to 2022.12. We use heteroskedasticity and autocorrelation consistent (HAC) standard errors of three lags, as formulated by Newey and West (1987), to compute p-values. Our findings show that the CNRI robustly predicts one-month to twelve-month commodity futures index excess returns, with coefficients ranging from 0.120 to 1.167 and adjusted R^2 values ranging from 0.83% to 3.68%. Notably, as the forecast horizon extends, both R^2 and β increase. Positive regression coefficients suggest that increases in the commodity news ratio index are generally followed by rises in subsequent commodity futures index excess returns.

Table 1

Commodity News Ratio Index and Commodity Futures Index Excess Return: In-Sample Results

	(1) 1946-2022		(2) 2004-2022	
Horizon	β	$\bar{R}^2$	β	$\bar{R}^2$
h=1	0.120	0.83	0.111	2.29
	(0.038)		(0.007)	
h=3	0.325	1.79	0.359	7.23
	(0.035)		(0.000)	
h=6	0.596	2.64	0.810	16.54
	(0.021)		(0.000)	
h=9	0.877	3.24	1.179	21.71
	(0.018)		(0.000)	
h=12	1.167	3.68	1.422	23.60
	(0.019)		(0.000)	

Notes: This table reports the parameter estimates and the adjusted R^2 values of the predictive regression: $R_{t+1:t+h} = \alpha + \beta CNRI_t + \epsilon_{t+1:t+h}$, h = 1, 3, 6, 9, 12, where $R_{t+1:t+h}$ is the accumulated commodity futures index excess returns from month $t+1$ and $t+h$, $CNRI_t$ is the aligned commodity news ratio index based on ChatGPT, and $\epsilon_{t+1:t+h}$ is an error term. In parentheses, we report the p-values calculated using three-lag heteroskedasticity and autocorrelation consistent (HAC) standard errors of Newey and West (1987), and represent significance level at 1%, 5%, and 10%. The sample periods are respectively from January 1946 to December 2022, and from January 2004 to December 2022, and the latter is generally regarded as the period of commodity financialization (*e.g.*, Da *et al*., 2024).

In recent years, financialization has emerged as a pivotal factor for understanding commodity market dynamics. Following Tang and Xiong (2012) and Da *et al*. (2024), we identify the year 2004 as a key turning point for commodity financialization. To evaluate the predictive performance of the CNRI, we also present results from in-sample predictive regressions analyzing commodity returns starting from 2004. The

findings indicate a significant enhancement in predictive results, with p-values spanning between 0.000 and 0.007 and adjusted R^2 varying from 2.29% to 23.60%. These results suggest that the CNRI demonstrates a stronger predictive capability during the period of commodity financialization.

In summary, the CNRI exhibits strong and positive significance in predicting commodity futures index excess returns.

Concluding Remarks

Accurate and reliable forecasts of commodity returns are essential for policymakers and investors. This paper develops a text-based commodity news ratio index (CNRI) using ChatGPT, derived from a comprehensive analysis of over 2.5 million articles sourced from nine newspapers. Our empirical results demonstrate that the ChatGPT-based CNRI is significantly effective in forecasting 1- to 12-month accumulated commodity futures index excess returns, as evidenced by both in-sample and out-of-sample regression analyses. Moreover, we control for business variables and economic indicators to examine the significance of predictability.

Our analysis further reveals that the CNRI exhibits enhanced predictive power during periods of economic expansion, contango markets, and declining inflation. We also confirm that our constructed CNRI contains valuable predictive information regarding future macroeconomic performance. In summary, our study illustrates that ChatGPT can effectively predict trends in the commodity market, enhancing our understanding of the information processing capabilities of large language models (LLMs) and their implications for investors in the financial market.

References

Baker, S. D., 2021, "The Financialization of Storable Commodities," *Management Science*, Vol. 67, No. 1, pp. 471-499.

Baker, S. R., Bloom, N. and S. J. Davis, 2016, "Measuring Economic Policy Uncertainty," *Quarterly Journal of Economics*, Vol. 131, No. 4, November, pp. 1593-1636.

Bybee, L., Kelly, B. T., Manela A. and D. Xiu, 2024, "Business News and Business Cycles," *Journal of Finance,* Vol. 79, No. 5, October, pp. 3105-3147.

Chen, J., Tang, G., Zhou, G. and W. Zhu, 2023, "ChatGPT and Deepseek: Can They Predict the Stock Market and Macroeconomy?" Available at SSRN: https://ssrn.com/abstract=4660148.

Chen, Y., Kelly, B. T. and D. Xiu, 2023, "Expected Returns and Large Language Models." Available at SSRN: https://ssrn.com/abstract=4416687.

Chen, Y.-C., Rogoff, K. S. and B. Rossi, 2010, "Can Exchange Rates Forecast Commodity Prices?", *Quarterly Journal of Economics,* Vol. 125, No. 3, August, pp. 1145-1194.

Da, Z., Tang, K., Tao, Y. and L. Yang, 2024, "Financialization and Commodity Markets Serial Dependence," *Management Science*, Vol. 70, No. 4, April, pp. 2122-2143.

Devlin, J., Chang, M., Lee, K. and K. Toutanova, 2018, "BERT: Pretraining of Deep Bidirectional Transformers for Language Understanding." Available at arXiv: https://arxiv.org/abs/1810.04805.

Fisher, A., Martineau, C. and J. Sheng, 2022, "Macroeconomic Attention and Announcement Risk Premia," *Review of Financial Studies*, Vol. 35, No. 11, November, pp. 5057-5093.

Fuertes, A.-M., Miffre, J. and G. Rallis, 2010, "Tactical Allocation in Commodity Futures Markets: Combining Momentum and Term Structure Signals," *Journal of Banking and Finance*, Vol. 34, No. 10, October, pp. 2530-2548.

Gargano, A. and A. Timmermann, 2012, "Predictive Dynamics in Commodity Prices," University of California and CREATES, San Diego, U.S.A. and Denmark, Manuscript.

Han, Y., Zhou, G. and Y. Zhu, 2016, "A Trend Factor: Any Economic Gains from Using Information over Investment Horizons?", *Journal of Financial Economics*, Vol. 122, No. 2, November, pp. 352-375.

Henderson, B. J., Pearson, N. D. and L. Wang, 2015, "New Evidence on the Financialization of Commodity Markets," *Review of Financial Studies*, Vol. 28, No. 5, May, pp. 1285-1311.

Hong, Y., Jiang, F., Meng, L. and B. Xue, 2024, "Forecasting Inflation with Economic Narratives and Machine Learning," *Journal of Business and Economic Statistics*, Vol. 43, No. 1, pp. 216-231.

Huang, D., Jiang, F., Tu, J. and G. Zhou, 2015, "Investor Sentiment Aligned: A Powerful Predictor of Stock Returns," *Review of Financial Studies*, Vol. 28, No. 3, March, pp. 791-837.

Kelly, B. and S. Pruitt, 2013, "Market Expectations in the Cross-Section of Present Values," *Journal of Finance*, Vol. 68, No. 5, October, pp. 1721-1756.

Kilian, L. and X. Zhou, 2022, "Oil Prices, Gasoline Prices, and Inflation Expectations," *Journal of Applied Econometrics*, Vol. 37, No. 5, August, pp. 867-881.

Levine, A., Ooi, Y. H., Richardson, M. and C. Sasseville, 2018, "Commodities for the Long Run," *Financial Analysts Journal*, Vol. 74, No. 2, December, pp. 55-68.

Lutzenberger, F. T., 2014, "The Predictability of Aggregate Returns on Commodity Futures," *Review of Financial Economics*, Vol. 23, No. 3, September, pp. 120-130.

Malik, F. and S. Hammoudeh, 2007, "Shock and Volatility Transmission in the Oil, U.S. and Gulf Equity Markets," *International Review of Economics and Finance*, Vol. 16, No. 3, pp. 357-368.

Manela, A. and A. Moreira, 2017, "News Implied Volatility and Disaster Concerns," *Journal of Financial Economics*, Vol. 123, No. 1, January, pp. 137-162.

Mensi, W., Hammoudeh, S., Shahzad, S. and M. Shahbaz, 2016, "Modeling Systemic Risk and Dependence Structure Between Oil and Stock Markets Using a Variational Mode Decomposition-Based Copula Method," *Journal of Banking and Finance*, Vol. 75, February, pp. 258-279.

Newey, W. K. and K. D. West, 1987, "A Simple, Positive Semidefinite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, Vol. 55, No. 3, May, pp. 703-708.

Omura, A. and N. Todorova, 2019, "The Quantile Dependence of Commodity Futures Markets on News Sentiment," *Journal of Futures Markets*, Vol. 39, No. 7, July, pp. 818-837.

Ready, R., Roussanov, N. and C. Ward, 2017, "Commodity Trade and the Carry Trade: A Tale of Two Countries," *Journal of Finance*, Vol. 72, No. 6, December, pp. 2629-2684.

Rogers, A., Kovaleva, O. and A. Rumshisky, 2021, "A Primer in BERTology: What We Know About How Bert Works," *Transactions of the Association for Computational Linguistics*, Vol. 8, pp. 842-866.

Silvennoinen, A. and S. Thorp, 2013, "Financialization, Crisis and Commodity Correlation Dynamics," *Journal of International Financial Markets, Institutions and Money*, Vol. 24, April, pp. 42-65.

Smales, L. A., 2014, "News Sentiment in the Gold Futures Market," *Journal of Banking and Finance*, Vol. 49, pp. 275-286.

Tang, K. and W. Xiong, 2012, "Index Investment and the Financialization of Commodities," *Financial Analysts Journal*, Vol. 68, No. 6, pp. 54-74.

Tetlock, P. C., 2007, "Giving Content to Investor Sentiment: The Role of Media in the Stock Market," *Journal of Finance*, Vol. 62, Vol. 3, June, pp. 1139-1168.

Wang, C., 2001, "Investor Sentiment and Return Predictability in Agricultural Futures Markets," *Journal of Futures Markets*, Vol. 21, No. 10, October, pp. 929-952.

Welch, I. and A. Goval, 2008, "A Comprehensive Look at the Empirical Performance of Equity Premium Prediction," *Review of Financial Studies*, Vol. 21, No. 4, July, pp. 1455-1508.

Wold, H., 1966, "Estimation of Principal Components and Related Models by Iterative Least Squares," in P. R. Krishnajah (ed), Multivariate Analysis, New York: Academic Press, pp. 391-420.

Zhang, Q., 2021, "One Hundred Years of Rare Disaster Concerns and Commodity Prices," *Journal of Futures Markets*, Vol. 41, Vol. 12, December, pp. 1891-1915.

Zhao, F., Zhou, G. and X. Zhu, 2021, "Unspanned Global Macro Risks in Bond Returns," *Management Science*, Vol. 67, No. 12, December, pp. 7291-7951.

Keywords

Commodity return predictability, textual analysis, ChatGPT.

Author Biographies

SHEN GAO, Ph.D.
School of Computer Science and Engineering, University of Electronic Science and Technology of China

Shen Gao is a professor at the School of Computer Science and Engineering at the University of Electronic Science and Technology of China (UESTC). Prior to this, he served as an assistant professor at the School of Computer Science and Technology, Shandong University. He received his Ph.D. from the Wangxuan Institute of Computer Technology (WICT), Peking University. His research interests lie primarily in natural language processing, with a focus on pre-trained language models, large language model (LLM)-based agents, and tool learning.

SHIJIE WANG
School of Economics, Shandong University, China

Shijie Wang is a Ph.D. candidate at Shandong University, China.

YUANZHI WANG
School of Economics, Shandong University, China

Yuanzhi Wang is a Ph.D. candidate at Shandong University, China.

QUNZI ZHANG, Ph.D.
School of Economic, Shandong University, China

Qunzi Zhang is Professor of Finance at the Shandong University, China. She obtained her Ph.D. in Finance from the University of Lausanne and Swiss Finance Institute. Her main research interests are asset pricing, investment and portfolio allocation, and innovation. Her papers have been published in a variety of academic journals including *Journal of Financial Economics*, *Management Science*, *Journal of Financial and Quantitative Analysis*, among others.