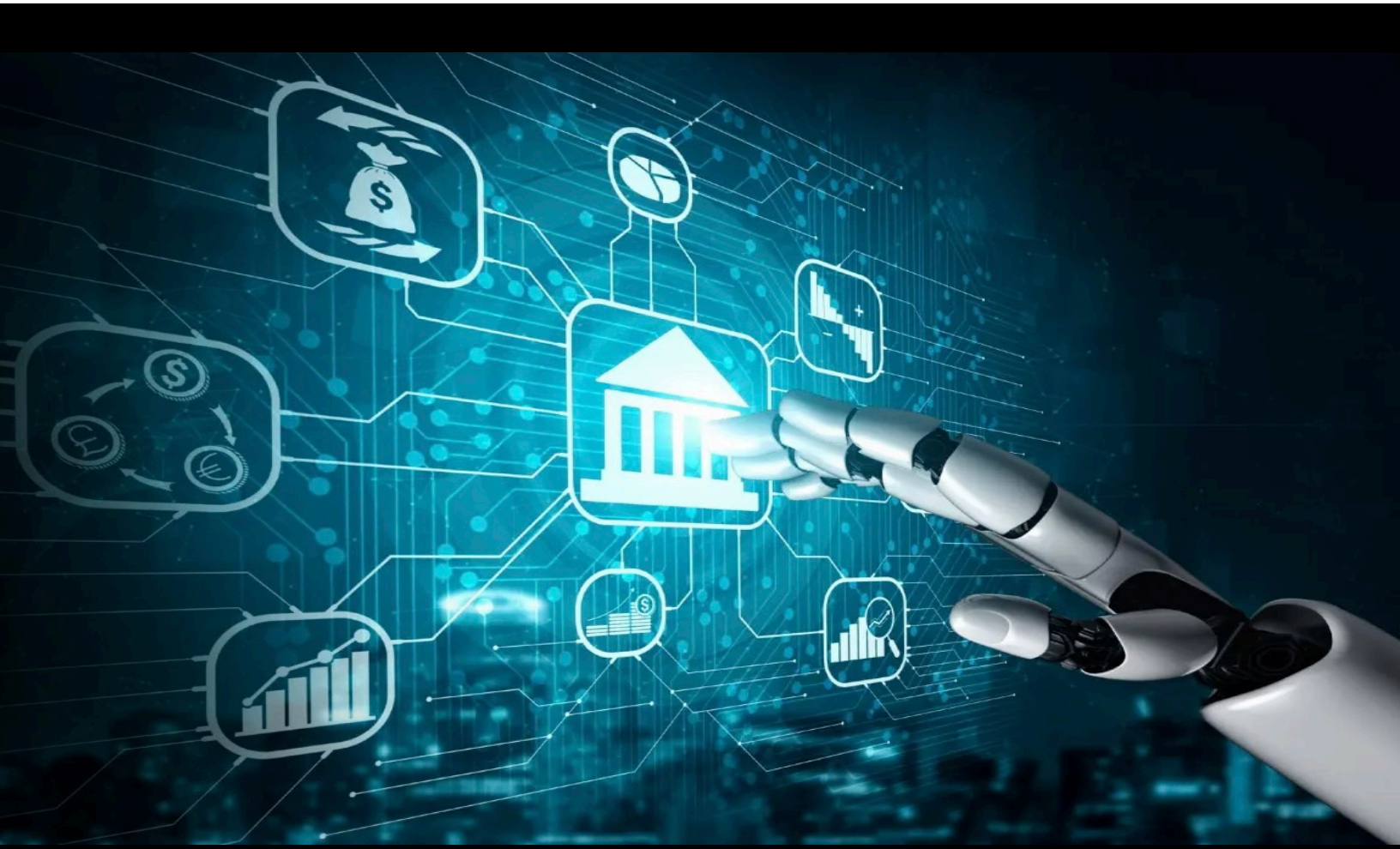
COMMODITY INSIGHTS

DIGEST

BAYES BUSINESS SCHOOL
CITY ST GEORGE'S UNIVERSITY OF LONDON

NOVEMBER 2025 - AI/ML IN COMMODITIES



Supported by funding from

PREMIA RESEARCH LLC

KONTANGO

COMMODITY INSIGHTS DIGEST

November 2025 Special Report on AI/ML in Commodities

Index of Articles

Research Digest Articles	Page
ChatGPT and Commodity Return *Shen Gao, Shijie Wang, Yuanzhi Wang, and Qunzi Zhang*	**4**
Predictability of Commodity Futures Returns with Machine Learning Models *Shirui Wang and Tianyang Zhang*	**12**
The Lead-lag Relations in the Commodity Futures Returns: A Machine Learning Approach *Yufeng Han and Lingfei Kong*	**17**
Media Emotion Intensity and Commodity Futures Pricing *Yeguang Chi, Lina El-Jahel, and Thanh Vu*	**23**
Forecasting Realized Volatility of Crude Oil Futures Prices Based on Machine Learning *Jiawen Luo, Tony Klein, Thomas Walther, and Qiang Ji*	**29**

Practitioner Insights	Page
From Theory to Practice: Myths and Mysteries About Speculation in the Oil Market *Ilia Bouchouev and Wu-Yen Sun*	**37**
AI Won't Take Our Job: It Will Reshape It *Victor Liu*	**48**
Mitigating AI Infrastructure Risks *Jodie Gunzberg and Hilary Till*	**56**

The *Commodity Insights Digest* (*CID*) is a publication of Bayes Business School, City St George's, University of London (U.K.).

The *CID* acknowledges the central importance and role of knowledge-exchange and engagement between academia and industry. Accordingly, the *CID*'s mission is to disseminate cutting-edge academic research along with industry insights that are based on hands-on experience in the commodity markets.

The themes of interest include commodity pricing and trading strategies, risk modeling and forecasting methods, as well as debates and challenges associated with climate change, geopolitical risks, and biodiversity.

The *CID* seeks to foster knowledge transfer between academics and practitioners by publishing: scholarly research digest articles; practitioner insights and interviews; articles on economic history; and book reviews.

Contributors to the *CID* include: (a) scholars seeking to enhance the impact of their research by delving into the practical implications of their theoretical or empirical studies; (b) consultants reporting on challenges faced by commodity market participants; and (c) industry economists providing analyses on their areas of market expertise.

The *CID* is a successor publication to the legacy *Global Commodities Applied Research Digest* (*GCARD*).

The *Commodity Insights Digest* is co-edited by Professor Ana-Maria Fuertes, Ph.D. in International Finance, and Honorary Senior Visiting Fellow, Hilary Till, M.Sc. in Statistics.

The *CID* benefits from the involvement of its distinguished Editorial Board. This international board consists of experts from across all commodity segments. The board is composed of academics, researchers, educators, policy advisors, and practitioners, all of whom have an interest in disseminating thoughtful research on commodities to a wider audience.

The corporate sponsors of the *CID* are Premia Research LLC and Kontango, Inc. In addition, the Payne Institute for Public Policy at the Colorado School of Mines has sponsored a Special Report on Critical Minerals.

Complimentary subscriptions to the *CID* are available at: https://www.bayes-cid.com/subscribe/.

Editorial Board

Managing Co-Editors
For Scholarly Research: ***Ana-Maria Fuertes, Bayes Business School, U.K.***
For Practitioner Insights: ***Hilary Till, Premia Research LLC, U.S.A.***

Amir ALIZADEH *Bayes Business School, U.K.*	**David JACKS** *Yale-NUS College, Singapore*
John BAFFES *World Bank, U.S.A.*	**Wouter JACOBS** *Erasmus Commodity & Trade Centre, Erasmus University Rotterdam, The Netherlands*
Keith BLACK *RIA Channel, U.S.A.*	**Hyun Jin JANG** *Ulsan National Institute of Science and Technology, South Korea*
Michael BOOTH *Hull Street Energy, U.S.A.*	**Daniel JERRETT** *SmartMacro Labs, U.S.A.*
Ilia BOUCHOUEV *Pentathlon Investments LLC, U.S.A.*	**Jan-Hein JESSE** *JOSCO Energy Finance & Advisory, The Netherlands*
Thomas BRADY *Colorado School of Mines, U.S.A.*	**Mark KEENAN** *Millenium, U.K.*
Bahattin BUYUKSAHIN *CoMeX Consulting and Advising, U.S.A.*	**John KOWALIK** *Derivative Applications, U.S.A.*
Kevin CHEN *Horizon Financial, U.S.A.*	**Dan KOWALSKI** *Land O'Lakes, Inc., U.S.A.*
Jennifer CONSIDINE *KAPSARC, Saudi Arabia*	**Alexander KUROV** *West Virginia University, U.S.A.*
Carlos Da COSTA *Canadian Institute of Mining, Metallurgy, and Petroleum, Canada*	**Thomas K. LEE** *U.S. Energy Information Administration*
Gillis BJÖRK DANIELSEN *Qube Research & Technologies, France*	**Yannick LE PEN** *Université Paris-Dauphine, France*
Charoula DASKALAKI *University of Crete, Greece*	**Julie LERNER** *PX Consultants, U.S.A.*
Rita D'ECCLESIA *Sapienza Università di Roma, Italy*	**Victor LIU** *Hemado Green Energy, U.K.*
Joseph EAGLEEYE *Premia Research LLC, U.S.A.*	**Michael MCGLONE** *Bloomberg Intelligence, U.S.A.*
John FAN *Griffiths Business School, Australia*	**Joëlle MIFFRE** *Audencia Business School, France*
Adrian FERNANDEZ-PEREZ *UCD Michael Smurfit Graduate Business School, University College Dublin, Ireland*	**Daniel MURRAY** *EFG Asset Management, Switzerland*
Jodie GUNZBERG *InFi Strategies, LLC, U.S.A.*	**Nikos NOMIKOS** *Bayes Business School, U.K.*
Fabian HOLLSTEIN *Universität des Saarlandes, Germany*	**Fergal O'CONNOR** *Cork University Business School, Ireland*

Editorial Board (Continued)

Peter O'NEILL *Archer Daniels Midland Company, U.S.A.*	**Marta SZYMANOWSKA** *Erasmus University Rotterdam, The Netherlands*
Marcel PROKOPCZUK *Leibniz Universität Hannover, Germany*	**Michael TAMVAKIS** *Bayes Business School, U.K.*
Xiao QIAO *Paraconic Technologies, U.S.A.*	**Lance TITUS** *GridStor, U.S.A.*
Deanna REITMAN *Faegre Drinker Biddle & Reath LLP, U.S.A.*	**Bruce TURNER** *Fonterra Co-op Group Ltd, New Zealand*
Kylie-Anne RICHARDS *Future Fund, Australia*	**Thomas WALTHER** *Utrecht University, The Netherlands*
Michel ROBE *University of Richmond, U.S.A.*	**Colin WAUGH** *C. Macdonald (Europe), U.K.*
Andrea RONCORONI *ESSEC Business School, France*	**Chardin WESE SIMEN** *University of Liverpool Management School, U.K.*
George SKIADOPOULOS *University of Piraeus, Greece*	**Sjur WESTGAARD** *Norwegian University of Science and Technology, Norway*
Sven STREITMAYER *Robert Bosch GmbH, Germany*	**Jian YANG** *University of Colorado Denver Business School, U.S.A.*
Lazaros SYMEONIDIS *Essex Business School, U.K.*	

TABLE OF CONTENTS
November 2025 Special Report on AI/ML in Commodities

Introduction

Introduction to the November 2025 Special Report on AI/ML in Commodities ii

Editorial Board

Editorial Board Listing iii

Research Digest Articles

ChatGPT and Commodity Return 4
By 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; and Qunzi Zhang, Ph.D., School of Economics, Shandong University, China

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, the authors demonstrate that their constructed commodity news ratio index significantly predicts future commodity returns, both in-sample and out-of-sample. Furthermore, it outperforms traditional textual analysis methods 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.

Predictability of Commodity Futures Returns with Machine Learning Models 12
By Shirui Wang, Ph.D., Independent Researcher, China; and Tianyang Zhang, Ph.D., Wenlan School of Business, Zhongnan University of Economics and Law, China

The authors use prevailing machine learning models to investigate the predictability of futures returns in 22 commodities with commodity-specific and macroeconomic factors as predictors. Out-of-sample prediction errors for the majority of futures contracts are lowered compared with those obtained by the baseline models of AR(1) and forecast combinations. Using Shapley values to explain feature importance, we identify dominant predictors for each commodity. A long–short portfolio strategy based on monthly Light Gradient-boosting Machine predictions outperforms the benchmark linear models in terms of annual return, Sharpe ratio, and maximum drawdown.

The Lead-lag Relations in the Commodity Futures Returns: A Machine Learning Approach 17
By Yufeng Han, Ph.D., Professor of Finance, Belk College of Business, University of North Carolina at Charlotte, U.S.A.; and Lingfei Kong, Ph.D., Assistant Teaching Professor of Finance, D'Amore - McKim School of Business, Northeastern University, U.S.A.

This paper uses machine learning tools to study the lead-lag relations in commodity futures returns. The authors use LASSO (Least Absolute Shrinkage and Selection Operator) to select the predictors because
(Continued on the next page.)

Research Digest Articles
(Continued)

the number of predictors is large relative to the number of observations. They find significant full-sample and out-of-sample predictability. The out-of-sample forecasts based on LASSO generate statistically and economically large gains. When the authors use more complex machine learning models such as regression trees and neural networks to forecast commodity futures returns, the out-of-sample performance is worse than LASSO portfolios, suggesting that nonlinearities and interactions do not appear substantial in the data. Finally, the authors find that index trading due to financialization has been associated with excess comovement among commodity futures.

Media Emotion Intensity and Commodity Futures Pricing **23**
By Yeguang Chi, Ph.D., University of Auckland, New Zealand; Lina El-Jahel, Ph.D., University of Auckland, New Zealand; and Thanh Vu, Ph.D., Auckland University of Technology, New Zealand

This study introduces media emotion intensity—measuring the proportion of emotional versus factual content in news—as a novel predictor of commodity futures returns. Emotion intensity measures the proportion of emotional content relative to factual content in media news. The media emotion intensity factor generates an economically significant annual premium of approximately 13%, even after accounting for transaction costs. This premium is more pronounced for commodities with low media coverage, high momentum, high basis-momentum, high hedging pressure, and backwardation. Emotion intensity significantly predicts the trading tendencies of both commercial and non-commercial traders and the cross-section of commodity futures returns at both portfolio and individual levels. The study also finds that media emotion intensity predicts future commodities' sentiment. Further, other commonly considered risk sources cannot subsume the predictability of the media emotion intensity factor.

Forecasting Realized Volatility of Crude Oil Futures Prices Based on Machine Learning **29**
By Jiawen Luo, Ph.D., School of Business Administration, South China University of Technology, China; Tony Klein, Ph.D., Faculty of Business and Economics, Technische Universität Chemnitz, Germany; Thomas Walther, Ph.D., Utrecht School of Economics, Utrecht University, The Netherlands and Editorial Board Member, Commodity Insights Digest; and Qiang Ji, Ph.D., Institutes of Science and Development, Chinese Academy of Sciences, China

Extending the popular heterogeneous autoregressive (HAR) model with additional information channels to forecast realized volatility of WTI futures prices, the authors show that machine learning-generated forecasts provide better forecasting quality and that portfolios that are constructed with these forecasts outperform their competing models resulting in economic gains. Analyzing the selection process, they show that information channels vary across forecasting horizons. Variable selection produces clusters and provides evidence that there are structural changes with regard to the significance of information channels.

Table of Contents

Practitioner Insights

From Theory to Practice: Myths and Mysteries About Speculation in the Oil Market **37**
By Ilia Bouchouev, Ph.D., Managing Partner, Pentathlon Investments LLC, U.S.A. and Editorial Board Member, Commodity Insights Digest; and Wu-Yen Sun, King Abdullah Petroleum Studies and Research Center (KAPSARC), KSA

The article provides unique quantitative insights about highly secretive and poorly understood speculation in the oil market. To demystify the behavior of speculators, the authors look at the problem from several different angles. First, they explain how the presence of the large over-the-counter (OTC) oil derivatives market leads to mischaracterization of traditional hedgers and speculators. They then explain what makes oil speculation special and different from speculation across many other commodities. Consequently, the authors identify the winners and the losers in the oil futures market. To model the behavior of the winners which they associate with fast-moving quantitative hedge funds, they develop a novel framework based on a simple neural-network algorithm. The authors conclude by analyzing a popular investment strategy of following the winners, or the hedge-funds' flows.

AI Won't Take Our Job: It Will Reshape It **48**
By Victor Liu, Ph.D., Managing Director, Hemado Green Energy, U.K. and Editorial Board Member, Commodity Insights Digest

This article notes that Artificial Intelligence (AI) is a tool, not a replacement. However, as it evolves, it will redefine our roles. Efficiency gaps will widen, the nature of work will change and standards of professionalism will rise. AI literacy is the way forward.

Mitigating AI Infrastructure Risks **56**
By Jodie Gunzberg, CFA, Managing Partner, InFi Strategies, U.S.A. and Editorial Board Member, Commodity Insights Digest; and Hilary Till, Managing Co-Editor, Commodity Insights Digest; Honorary Senior Visiting Fellow, Bayes Business School, City St George's, University of London, U.K.; and Principal, Chicago-based Premia Research LLC, U.S.A.

This article examines the risks associated with AI infrastructure investments, focusing on energy consumption, natural resource constraints, and historical parallels from past industrial buildouts. It explores factors that could lead to infrastructure obsolescence, market misalignment, and the ways that investment strategies may evolve in response to AI's future.

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.

Predictability of Commodity Futures Returns with Machine Learning Models

Shirui Wang, Ph.D.
Independent Researcher, China

Tianyang Zhang, Ph.D.
Wenlan School of Business, Zhongnan University of Economics and Law, China

Published in *Journal of Futures Markets* and Available at: https://doi.org/10.1002/fut.22471

We use prevailing machine learning models to investigate the predictability of futures returns in 22 commodities with commodity-specific and macroeconomic factors as predictors. Out-of-sample prediction errors for the majority of futures contracts are lowered compared with those obtained by the baseline models of AR(1) and forecast combinations. Using Shapley values to explain feature importance, we identify dominant predictors for each commodity. A long–short portfolio strategy based on monthly Light Gradient-boosting Machine predictions outperforms the benchmark linear models in terms of annual return, Sharpe ratio, and maximum drawdown.

Introduction

The predictability of futures returns is one of the main questions of interest in the commodity markets in recent decades. As an alternative asset class from traditional financial assets such as equities, the factors that price the traditional asset classes are not appropriate for commodities (Giampietro *et al*., 2018; Gorton and Rouwenhorst, 2006). Besides commodity-specific factors, commodity markets are also well known to be affected by macroeconomic factors. There is still debate about the predictability of risk factors since previous studies have shown mixed results. Examples include Daskalaki *et al*. (2014), Fernandez-Perez *et al*. (2018), and others.

Most studies in the commodity markets mainly use the traditional asset pricing models, such as linear regression models. However, the underlying data-generating process of asset returns is likely nonlinear, thus predictions are typically improved by adding nonlinearities while accounting for overfitting. We employ machine learning methods that allow for nonlinear forms of predictors that are not captured by traditional linear pricing models.

Our paper has several contributions. First, we apply modern machine learning techniques to select and interpret factors, and find that the significant predictive factors of each commodity are distinct, which explains why there is no agreement on the common predictive factors in the commodity markets based on linear relationships. This indicates that it is critical to focus on each contract instead of doing a cross-sectional test as in most previous research in the literature. Second, we implement a comprehensive comparison of models and contracts, and find that Light Gradient-boosting Machine (LightGBM) performs the best in predicting futures returns. Third, we construct a practical long–short portfolio strategy based on return predictions that outperform strategies with AR(1) and forecast combinations.

Data and Methodology

The data include futures contracts for 22 commodities, covering the period from May 1997 to December 2018. For each commodity, we construct continuous time series monthly futures returns. Following Gu *et al*. (2020) and Leippold *et al*. (2022), we adopt an additive prediction error model to study the predictability of futures return of commodity i at time $t+1$ ($r_{i,t+1}$) when all available risk factors are included:

$$r_{i,t+1} = f(Z_{i,t}) + \varepsilon_{i,t+1} \quad (1)$$

where $Z_{i,t}$ is a vector of predictors at time t. The specific formula of $f(\cdot)$ is undefined and dependent on the model of choice.

The predictors of futures return are sorted into two groups: commodity-specific and macroeconomic factors. Commodity-specific factors are constructed for each commodity, while the macroeconomic factors are common for all commodities. On the basis of the findings in the previous studies, 15 commodity-specific factors used in this study are lagged return, liquidity factor, futures basis, hedging pressure, moving average of hedging pressure, call-to-put volume ratio, skewness, past 6-month futures returns, past 12-month futures returns, the difference between lagged return and past 6-month futures returns, the difference between lagged return and past 12-month futures returns, the logarithm value of futures open interest, the logarithm value of futures trading volume, the average lagged return in the sector, and the average lagged return in the whole market. To capture seasonal information, we calculate the rolling mean, standard deviation, minimum, and maximum in the same month with past values for the commodity-specific factors. In addition, we construct the interaction terms that are formed by two commodity-specific factors. To control for the effect of market volatility, we include the squared of market return. As a result, the total number of factors that are related to commodity markets is 181.

Monthly macroeconomic factors are combined from the updated data of Welch and Goyal (2008) and McCracken and Ng (2016). If there are common factors in these two databases, we only keep the ones from the database in McCracken and Ng (2016). As a result, there are 136 macroeconomic factors included in our study. In the appendix of the published paper, we provide the explanations of the factors selected by the model for prediction. Details of other macroeconomic factors are referred to Welch and Goyal (2008) and McCracken and Ng (2016). We also include three seasonal dummy variables. Thus, the total number of risk factors in this study is 320.

We choose five prevailing machine learning models: Elastic net (ENET), principal component regression (PCR), support vector regression (SVR), LightGBM (LGB), and Neural Network (NN). Modern machine learning methods are particularly good at feature selection and extraction of nonlinear relationships among features. To this end, we follow the standard practitioner's procedure by first creating statistics and interactions terms to create a rich set of features, and then allow models to select features based on prediction errors of validation sets. Feature selection is essential to balance the bias-variance trade-off and avoid overfitting.

Prediction Results

In this study, we employ AR(1) and forecast combinations models as two benchmark models. The out-of-sample mean squared predictions error (MSE) was used to assess the predictive performance of AR(1) and forecast combinations models, along with five machine learning models. For commodity i, this is defined as

$$MSE(i) = \frac{1}{N}\sum_{(i,t)\in\Theta}(r_{i,t} - \hat{r}_{i,t+1}) \quad (2)$$

where $r_{i,t}$ and $\hat{r}_{i,t+1}$ are observed and predicted futures return, N is the number of observations in the out-of-sample test and Θ indicates the fits that are only assessed on the testing subsample. For each commodity, we make a 70%/20%/10% time split for the training, validation, and test sets.

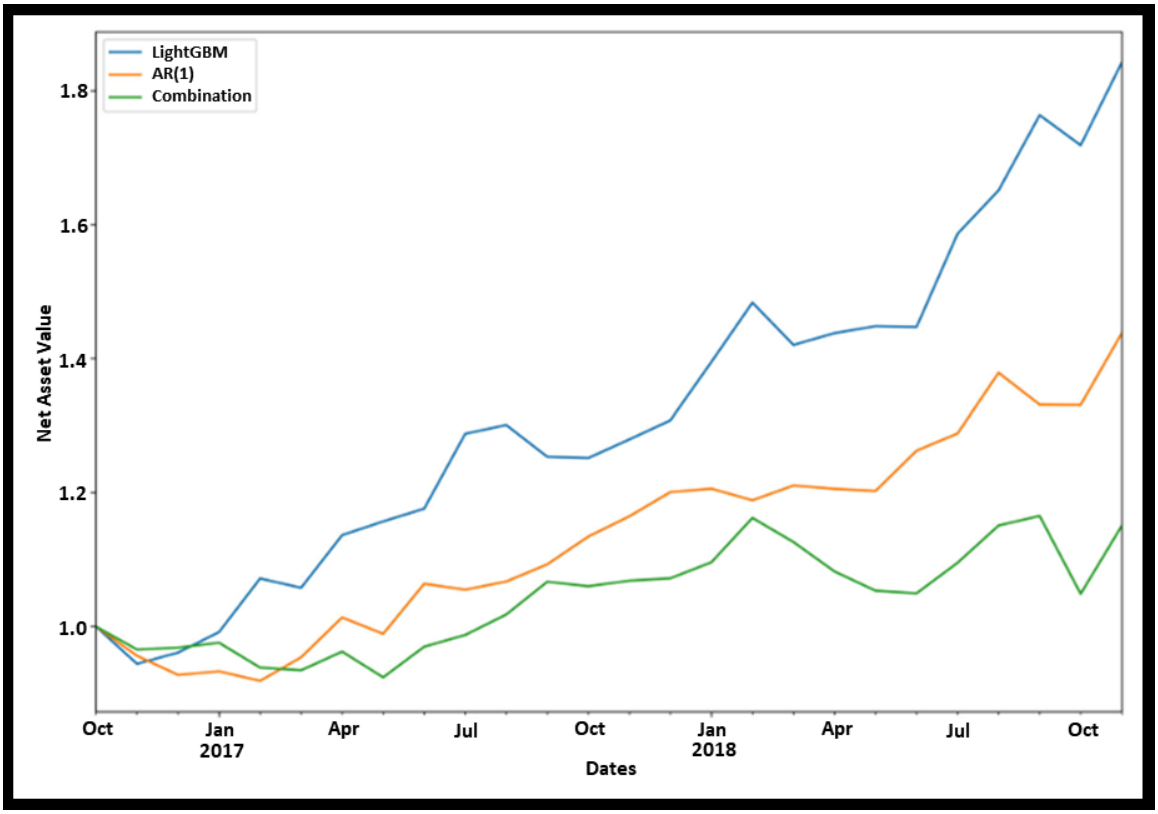
All predictive variables are included when comparing different models, although which factors are fed into the final models depends on whether they survive feature selections that are distinct for each model. The results show that the lowest out-of-sample MSE values differ a lot across different commodities. For the average MSEs among all commodities, LightGBM is the lowest with 0.438, followed by AR(1), ENET, NN, PCR, forecast combinations, and SVR. The results of the CW-test also show that LightGBM has significantly better predictive performance than AR(1) and forecast combinations.

In addition to analyzing the out-of-sample MSE values of each model, we use Shapley values obtained from LightGBM to quantify the importance of each factor in affecting futures returns and rank all factors based on the predictive performance for each commodity. All commodities are categorized into four broad sectors: agriculture (13 commodities), energy (three commodities), livestock (three commodities), and metals (three commodities). In summary, there are no common factors in predicting futures returns since the top performance factors are distinct for each commodity.

The empirical results show that the factors related to hedging pressure and momentum perform best among all the factors associated with commodity markets. The fact that the hedging pressure effects only exist in some commodities is consistent with the finding that there is not an agreement on the hedging pressure effect in commodity markets. Regarding macroeconomic factors, the information mainly from labor markets, industrial production, credit markets, interest rate markets, and stock markets has predictive power for certain commodities.

Next, we take advantage of the machine learning methods in predicting expected returns and evaluate trading strategies based on the predictions. A market-neutral strategy is built by going long the contracts with the highest predicted returns and shorting the losers. Figure 1 on the next page shows the net asset values of the three strategies. The LightGBM strategy achieves a 35% annualized return and 2.07 Sharpe ratio in the testing sample, compared to the 20% (8%) and 1.52 (0.54) Sharpe ratio of the AR(1) (forecast combinations) strategy in the same period. The LightGBM strategy's performance is consistently best over time, except for the period in the first two months of the testing sample.

Figure 1
Net Asset Values of LightGBM, AR(1), and Forecast Combination Strategies, October 2016-November 2018



Conclusion

This study investigates the performance of several machine learning models in predicting commodity futures returns for 22 commodities. All the risk factors are sorted into two groups that are commodity-specific and macroeconomic factors. Compared with benchmark AR(1) and forecast combinations models, LightGBM performs the best among other machine learning models in terms of out-of-sample MSE overall. The top predictive factors based on the Shapley values are provided for each commodity. Interestingly, we do not find common predictors for all commodities, which indicates the importance of analyzing commodities individually. Overall, the performance of factors that are associated with momentum and hedging pressure are better than the other factors related to commodity markets. In addition, the macroeconomic factors demonstrating predictive power for futures returns belong to multiple groups, with labor and stock markets serving as two examples. A long–short strategy based on predicted returns with LightGBM is shown to be profitable and superior to the benchmark, with a 35% annualized return and a Sharpe ratio of 2.07. The downside risk control is also practically desirable.

Endnote

All errors are our own. The authors declare no conflict of interest.

References

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

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

Giampietro, M., Guidolin, M. and M. Pedio, 2018, "Estimating Stochastic Discount Factor Models with Hidden Regimes: Applications to Commodity Pricing," *European Journal of Operational Research*, Vol. 265, No. 2, March, pp. 685-702.

Gorton, G. and K. G. Rouwenhorst, 2006, "Facts and Fantasies about Commodity Futures," *Financial Analysts Journal*, Vol. 62, No. 2, March-April, pp. 47-68.

Gu, S., Kelly, B. and D. Xiu, 2020, "Empirical Asset Pricing via Machine Learning," *The Review of Financial Studies*, Vol. 33, No. 5, May, pp. 2223-2273.

Leippold, M., Wang, Q. and W. Zhou, 2022, "Machine Learning in the Chinese Stock Market," *Journal of Financial Economics*, Vol. 145, No. 2, August, pp. 64-82.

McCracken, M. W. and S. Ng, 2016, "FRED-MD: A Monthly Database for Macroeconomic Research," *Journal of Business & Economic Statistics*, Vol. 34, No. 4, October, pp. 57-589.

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

Keywords

Commodity futures, machine learning, LightGBM.

Author Biographies

SHIRUI WANG, Ph.D.
Independent Researcher, China

Shirui Wang is an independent researcher and crypto trader. He received his Ph.D. in Economics from Iowa State University in 2020. His research focuses on commodity and crypto markets. He has published papers in *Journal of Futures Markets,* and *Economic Theory and Business Management (Chinese).*

TIANYANG ZHANG, Ph.D.
Wenlan School of Business, Zhongnan University of Economics and Law, China

Tianyang Zhang is an Assistant Professor in the Wenlan School of Business, Zhongnan University of Economics and Law. He received his Ph.D. in Economics from Iowa State University in 2020. His fields of interest are within commodity markets and cryptocurrency markets. He has published papers in *Journal of Futures Markets*, *The North American Journal of Economics and Finance*, and *Finance Research Letters*.

The Lead-lag Relations in the Commodity Futures Returns: A Machine Learning Approach

Yufeng Han, Ph.D.
Belk College of Business, University of North Carolina at Charlotte, U.S.A.

Lingfei Kong, Ph.D.
D'Amore - McKim School of Business, Northeastern University, U.S.A.

Available at SSRN: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3536046

This paper uses machine learning tools to study the lead-lag relations in commodity futures returns. We use LASSO (Least Absolute Shrinkage and Selection Operator) to select the predictors because the number of predictors is large relative to the number of observations. We find significant full-sample and out-of-sample predictability. In the full sample, we find that LASSO can identify a sparse set of predictors that either come from economically linked commodities or from factors that are coincident with the period of "financialization" in the commodity markets. The out-of-sample forecasts based on LASSO generate statistically and economically large gains. When we use more complex machine learning models such as regression trees and neural networks to forecast commodity futures returns, the out-of-sample performance is worse than LASSO portfolios, suggesting that nonlinearities and interactions do not appear substantial in the data. Finally, we find that index trading due to financialization has been associated with excess comovement among commodity futures.

Introduction

Commodity futures contracts are agreements to buy or sell a predetermined quantity of a commodity at a specified price on a particular date in the future. Historically, commodity futures were primarily used by farmers, producers, and commodity merchandisers to lock in the price and reduce the risk of financial losses from price changes. Over the past two decades, financial institutions such as hedge funds, swap dealers, and mutual funds have dramatically increased their exposure to commodities. The advent of commodity futures Exchange-Traded Funds (ETFs) and Exchange-Traded Notes (ETNs) also gives individual investors easy access to commodity futures. According to BarclayHedge, the assets under management for managed futures have grown from $95.7 billion to $318.4 billion from 2003 to 2019; managed futures programs, in turn, include the trading of both financial futures and physical commodity futures (Gupta and Wilkens, 2007). The phenomenon of investor participation in physical commodity futures trading is called "financialization" by researchers (*e.g.*, Tang and Xiong, 2012; Basak and Pavlova, 2016), and has received extensive attention from researchers and practitioners.

This paper examines the predictability of commodity futures returns. Specifically, we examine the lead-lag relations in the commodity futures returns by predicting commodity futures returns with the commodity futures' own lagged returns as well as other commodity futures' lagged returns. We focus on the post-financialization period (2004-2019) when commodity futures became a popular investment asset class, and futures markets became more liquid than before. The inclusion of the broad set of lagged returns is motivated by the following arguments. First, many futures have economic links that can drive the lead-lag relations among futures returns. For instance, (a) heating oil is refined from crude oil so via the crack spreads, the lagged returns of crude oil futures would be tied to the returns of heating oil futures; (b) corn and soybean meal are the primary feeds for livestock such as pigs and cows, thus via the feed margins, the returns of corn and soybean futures would plausibly impact the returns of lean hogs and live

cattle futures at some stage of the production; and (c) crops such as corn and soybeans are used to produce biofuels, which are substitutes for fossil fuels such as crude oil, so crude oil futures prices would be expected to be related to grain futures prices. Mensi *et al*. (2014) find significant spillovers among grain and energy commodity prices by studying the lead-lag relations of the returns and volatilities of these commodities. Second, the increased trading activities by financial institutions and retail investors can either weaken or strengthen these effects, or even introduce new relations that are not driven by any economic links; for example, Till (2008) describes how seemingly unrelated commodities and financial instruments became (temporarily) very correlated during "risk off" episodes. Unlike producers/merchants who use futures contracts to lock in the price of their products or inputs, these traders generally do not have physical exposure to the commodities but use commodity futures as investment tools, and thus are non-commercial traders.

Some researchers have concluded that the proliferation of non-commercial trading may have caused price distortion and excessive price comovement, depending on the frequency of observation. For example, Tang and Xiong (2012) find that since 2004 when financial institutions significantly increased their positions on commodity futures, the prices of non-energy commodity futures have become increasingly correlated with oil futures, and the correlations are stronger for commodities that are included in the broadly traded commodity indices than for non-indexed commodities.

To provide an even balance of views, one must also note that there may be other explanations besides the proliferation of non-commercial futures traders for the increased correlations of oil futures with non-energy commodity futures, which would include a broad-based increase in commodity demand from emerging economies (Cevik and Sedik, 2011). Further, IEA (2011) states that "[a] … comparison of the non-exchange-traded commodity price index, as well as [the] crude oil price series, supports the notion that, starting in 2003 and more strongly after 2004, a demand shock pushed upward the price of most commodities." In addition, based on an examination of crude oil, natural gas, and corn futures markets, Brunetti *et al.* (2015) "find little evidence that speculators destabilize … markets."

Relevance of the Research Question

This paper contributes to the existing literature in the following ways. First, the majority of the literature, after financialization, studies the comovement among commodity futures in the context of contemporaneous relations rather than lead-lag relations. Moreover, most papers regarding return predictability in commodity futures focus on the futures' own past returns or other characteristics but ignore the cross-serial correlations. As far as we can tell, this paper is the first to analyze the predictability of the lagged futures returns that allows each individual commodity futures' return to respond to the lagged returns for all commodity futures, thereby accommodating a large dimension of commodity links, both direct and indirect. Second, we use machine learning techniques to overcome the potential overfitting problem. With a large number of predictors, OLS (ordinary least squares) estimation is subject to overfitting. Third, we find that incorporating the lagged returns of commodity futures can help to forecast the returns for individual commodity futures and construct profitable trading strategies. The performance of the long-only futures indices has been lackluster in the last decade. By contrast, the

actively constructed single-sort or timing strategies based on LASSO generate both statistically and economically large out-of-sample returns.

Data and Methodology

We study 27 commodity futures that are traded actively in the U.S. The data are collected from Bloomberg. The 27 commodity futures cover five main sectors, namely, grains, softs, livestock, energy, and metals. There are 8 grains futures, 7 softs futures, 3 livestock futures, 4 energy futures, and 5 metals futures in the sample. The sample period is from January 2004 to December 2019.

The main methodology used to forecast commodity futures returns is LASSO (Least Absolute Shrinkage and Selection Operator), as proposed by Tibshirani (1996). The reason why we use LASSO is that the number of predictors is large compared with the number of observations used to fit the model. With a large number of predictors, OLS model causes overfitting, which means it reduces the in-sample mean squared error but often results in a large out-of-sample mean squared error. LASSO selects a subset of predictors and thus reduces overfitting. We use the corrected Akaike information criterion (AICC) to select the shrinkage parameter in LASSO. We use OLS to estimate the coefficients for the predictors selected by LASSO (OLS post-LASSO) since LASSO can overshrink the magnitudes of the coefficients for the selected predictors (Efron *et al*., 2004). Our main benchmark forecast models are OLS that include all the lagged returns as predictors and prevailing mean, which is the moving average of the commodity futures returns during the estimation window.

In the in-sample analysis, we use the full sample to estimate the OLS post-LASSO regression. In the out-of-sample analysis, we run a 60-month rolling LASSO regression and calculate the one-step ahead out-of-sample return forecast. To evaluate the out-of-sample forecast performance, we construct two trading strategies, one is a single-sort portfolio, in which we form equally-weighted portfolios using the five commodities with the highest forecasted returns and the five commodity futures with the lowest forecasted returns. We then calculate the return differences between the two portfolios. The other strategy is a moving-average timing portfolio; that is, we establish a long position if its forecasted return is positive and a short position if its forecasted return is negative, then form an equally-weighted portfolio of all the positions. There are two versions of LASSO models in the out-of-sample analyses: they are LASSO that excludes any futures that do not have any lagged returns selected by the LASSO, and LASSO (All) that includes all contracts including those without any predictors.

To test whether financialization has been associated with the excess comovement among commodity futures, we have both in-sample tests and out-of-sample tests. In the in-sample test, we create a dummy variable that is equal to 1 when the commodity futures are members of indices and when the date is after the inception of three ETF or ETN products (September 2006) that track the returns of the Bloomberg Commodity Index (BCOM) or the S&P Goldman Sachs Commodity Index (GSCI); otherwise, the dummy variable is set to 0. We then add the dummy as well as its interaction with all the lagged returns as the additional predictors and run the OLS post-LASSO regression using the full sample. In the out-of-sample test, we classify the commodity futures included in both the BCOM and GSCI as indexed futures, and those included neither in the BCOM nor GSCI index as non-indexed futures. There are 16 indexed futures and 7

non-indexed futures. We then reconstruct the single sort and timing portfolios using the 16 indexed futures or the 7 non-indexed futures. We also test the lead-lag relations and construct single-sort and timing portfolios using data during 1970-2003, which is before the financialization period.

Main Results

The LASSO single-sort portfolio has an annualized mean return of 15.15%, which is more than eight times the mean return of the prevailing mean portfolio (1.78%). The annualized Sharpe ratio of this LASSO portfolio is also much higher than those of the two benchmark portfolios, 0.93 versus 0.10 (prevailing mean) and -0.21 (OLS). The LASSO timing portfolio yields an average excess return of 6.15% per year and a Sharpe ratio of 0.72. In contrast, the prevailing mean timing portfolio generates an annual return of only 1.80%, and the OLS timing portfolio generates an even smaller average return that is less than 1%. Overall, the results show that lagged returns contain predictability for the cross-section of commodity futures returns and LASSO helps to identify important predictors. The above results are robust to transaction costs, an alternative number of lagged returns, seasonality, and different machine learning models such as elastic net.

When we interact the ETF inception dummy with lagged returns for the whole sample, at least one lagged commodity futures return is selected for 14 of the 16 indexed futures. In addition, 10 out of the 14 futures have interactions of the ETF inception dummy with lagged returns as selected effects, and the hypothesis that the interactions are all equal to 0 is rejected for all the 10 futures. When we use indexed-only futures to run the lead-lag predictive regressions and construct portfolios, the LASSO portfolio generates economically and statistically significant out-of-sample returns in both the single-sort and timing strategies. However, when we use the non-indexed futures to run the lead-lag predictive regressions, the LASSO portfolio no longer generates any significant out-of-sample returns. In addition, when we examine the lead lag relations from 1970-2003, the two LASSO portfolios do not generate significant out-of-sample returns.

Conclusions

In this paper, we exploit the lead-lag relations in the commodity futures returns via machine learning tools (mainly LASSO) to analyze return predictability in commodity futures markets. The portfolio analysis reveals that the LASSO strategies outperform the well-known OLS and historical moving average by generating statistically and economically larger excess returns. We find that LASSO strategies perform well after the financialization period but not before financialization. When we separate the indexed futures from the non-indexed futures, we find that LASSO strategies only work in the basket of indexed futures. Moreover, the lead-lag relations are stronger after the advent of ETFs or ETNs that track the broad futures indices such as the GSCI and BCOM indices. Overall, our results suggest that lagged returns contain valuable information to predict futures returns and that machine learning is a useful and effective tool to increase out-of-sample predictability.

Endnote

Dr. Kong presented on this topic at the J.P. Morgan Center for Commodities' 5th Annual International Symposium at the University of Colorado Denver Business School on August 15, 2022.

References

Basak, S. and A. Pavlova, 2016, "A Model of Financialization of Commodities," *The Journal of Finance*, Vol. 71, No. 4, August, pp. 1511-1556.

Brunetti, C., Büyükşahin, B. and J. Harris, 2015, "Speculators, Prices and Market Volatility, Bank of Canada Working Paper, November.

Cevik, S. and T. Sedik, 2011, "A Barrel of Oil or a Bottle of Wine: How Do Global Growth Dynamics Affect Commodity Prices?", International Monetary Fund Working Paper, January.

Efron, B., Hastie, T., Johnstone, I. and R. Tibshirani, 2004, "Least Angle Regression," *The Annals of Statistics*, Vol. 32, No. 2, April, pp. 407-499.

Gupta, B. and K. Wilkens, 2007, "An Evaluation of the *Commodity* Exposure in Commodity Trading Advisers," in H. Till and J. Eagleeye (eds) Intelligent Commodity Investing, London: Risk Books, pp. 179-186.

IEA [International Energy Agency], 2011, "Oil Market Report," March 15, pp. 44-45.

Mensi, W., Hammoudeh, S., Nguyen, D. K. and S. M. Yoon, 2014, "Dynamic Spillovers Among Major Energy and Cereal Commodity Prices," *Energy Economics*, Vol. 43, May, pp. 225-243.

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

Tibshirani, R., 1996, "Regression Shrinkage and Selection Via the Lasso," *Journal of the Royal Statistical Society: Series B (Methodological)*, Vol. 58, No. 1, pp. 267-288.

Till, H., 2008, "Case Studies and Risk Management in Commodity Derivatives Trading," in H. Geman (ed) Risk Management in Commodity Markets, Chichester (U.K.): John Wiley & Sons Ltd., pp. 255-291.

Keywords

Commodity futures, LASSO, machine learning, predictability, lead-lag relations, financialization, comovement.

Author Biographies

YUFENG HAN, Ph.D.
Belk College of Business, University of North Carolina at Charlotte, U.S.A.

Dr. Yufeng Han joined the Belk College faculty as an Associate Professor of Finance in 2016. Prior to joining UNC Charlotte, he served as a faculty member at the University of Colorado Denver and Tulane University. Dr. Han's primary research interests are empirical asset pricing, investment, mutual funds, and econometrics. He has published in the *Review of Financial Studies, Journal of Financial Economics, Journal of Financial and Quantitative Analysis, Journal of Banking & Finance, Real Estate Economics* and many other finance and economics journals. Dr. Han received a Bachelor's degree from Tsinghua University (China) and a Ph.D. from Washington University in St. Louis.

The Lead-lag Relations in the Commodity Futures Returns: A Machine Learning Approach

LINGFEI KONG, Ph.D.
D'Amore - McKim School of Business, Northeastern University, U.S.A.

Dr. Lingfei Kong is an Assistant Teaching Professor of Finance at D'Amore-McKim School of Business at Northeastern University. Prior to joining Northeastern University, she was a postdoctoral research associate at Washington University in St Louis. She holds a Ph.D. in Finance and is a certified Financial Risk Manager (FRM). Her teaching and research interests focus on investments, data analytics, derivatives, and corporate finance.

Media Emotion Intensity and Commodity Futures Pricing

Yeguang Chi, Ph.D.
University of Auckland, New Zealand

Lina El-Jahel, Ph.D.
University of Auckland, New Zealand

Thanh Vu, Ph.D.
Auckland University of Technology, New Zealand

Published in *Journal of Commodity Markets* and Available at: https://doi.org/10.1016/j.jcomm.2025.100460

This study introduces media emotion intensity—measuring the proportion of emotional versus factual content in news—as a novel predictor of commodity futures returns. Emotion intensity measures the proportion of emotional content relative to factual content in media news. The media emotion intensity factor generates an economically significant annual premium of approximately 13%, even after accounting for transaction costs. This premium is more pronounced for commodities with low media coverage, high momentum, high basis-momentum, high hedging pressure, and backwardation. Emotion intensity significantly predicts the trading tendencies of both commercial and non-commercial traders and the cross-section of commodity futures returns at both portfolio and individual levels. The study also finds that media emotion intensity predicts future commodities' sentiment. Further, other commonly considered risk sources cannot subsume the predictability of the media emotion intensity factor.

Introduction

In recent decades, the financialization of commodity markets has intensified the search for factors that can explain and predict the cross-section of futures returns (Tang and Xiong, 2012). While the equity market has a vast library of predictive signals, the number of established factors in commodity markets remains limited, with benchmarks such as basis, momentum, and basis-momentum being the most prominent (Szymanowska *et al*., 2014; Boons and Prado, 2019).

The role of media in asset pricing has been well-documented, primarily through two lenses: news sentiment (the balance of positive versus negative words) and media coverage (the volume of news) (Tetlock, 2007; Fang and Peress, 2009). This study extends existing literature by introducing a novel dimension of media influence: media emotion intensity, defined as the proportion of emotional versus factual content in news. Rather than measuring *if* news is positive or negative, our factor measures the *degree* to which news content is emotional versus factual. For instance, even if the overall sentiment of two articles about oil is neutral, one might be written in a dry, data-driven style while the other uses evocative and emotional language to describe the same events. It is this crucial distinction that our factor captures.

The theoretical motivation for the media emotion intensity factor is twofold. First, the media strategically employs emotion to attract attention and engage readers, a trend that has accelerated in the digital age (Beckett, 2015). Second, a rich body of neuroscientific and psychological research shows that humans struggle to ignore emotion in decision-making, which can trigger behavioral biases and override rational

processes, especially under uncertainty (Vuilleumier, 2005; Loewenstein, 1996). Our research proposes that high emotional intensity in media can trigger behavioral biases among investors, leading to systematic and predictable patterns in commodity prices. We demonstrate that this factor is distinct from news sentiment (with low cross-sectional and time-series correlations) and offers significant, new explanatory power for understanding commodity futures returns.

Data and Methodology

Data and Factor Construction

This study uses futures data for 26 commodities from January 1998 to February 2020 from Barchart (formerly CRB). The key variable, media emotion intensity, is constructed using the Thomson Reuters MarketPsych (TRMI) database, which analyzes content from thousands of financial news outlets in real-time. This database calculates an "EmotionVsFact" index for each asset, quantifying the excess proportion of emotional over factual content, scaled by the total amount of media coverage ("buzz"). A higher value of this index indicates a greater intensity of emotional content relative to factual content in the news for a specific commodity. The benchmark factors include basis, momentum, and basis-momentum, following the Boons and Prado (2019) methodology.

Portfolio Strategy

The core of the methodology involves a long-short portfolio strategy. At the end of each month, 26 commodities are ranked based on their media emotion intensity value from the preceding month. A portfolio is then formed by taking an equal-weighted long position in the four commodities with the highest intensity ("High4") and an equal-weighted short position in the four with the lowest intensity ("Low4"). The performance of this "High4-Low4" portfolio is tracked over the following month.

Empirical Results

The key finding is that media emotion intensity is a powerful and robust predictor of future commodity returns. Table 1 on the next page reports the performance of the High4-Low4 portfolio strategy and compares it to portfolios built on other established factors. The media emotion intensity portfolio delivers a statistically and economically significant annualized return of 13.51% (t-statistic = 2.89), outperforming established benchmarks such as momentum (9.57%) and basis-momentum (12.50%). The premium of media emotion intensity survives a set of robustness checks:

- Transaction Costs: The net annualized return after transaction costs is 13.00% (t-statistic = 2.79).
- Portfolio Construction: The effect is persistent when using different portfolio sizes (*e.g.*, High6-Low6, High8-Low8).
- Spanning Tests: In time-series spanning tests, the "alpha" (abnormal return) of the emotion intensity portfolio remains statistically significant at around 10%, even when benchmarked against

a model including six key factors (average, basis, momentum, basis-momentum, hedging pressure, and value). The low R-squared values in these tests (around 0.34) confirm that the premium is not explained by existing benchmarks.

In contrast, portfolios based solely on traditional news sentiment fail to generate statistically significant returns, underscoring the unique and incremental informational value provided by the emotion intensity factor.

Table 1

Performance of Long-Short Portfolios Sorted by Media and Benchmark Factors (Nearby Returns, 1998-2020)

Sorting Factor	Portfolio	Ave. Ret.	t-statistic
Emotion Intensity	**High4 - Low4**	**13.51%**	**2.89**
Basis-momentum	High4 - Low4	12.50%	2.66
Momentum	High4 - Low4	9.57%	1.76
Media Coverage	Low4 - High4	9.16%	2.76
News Sentiment	High4 - Low4	5.60%	1.11
Basis	Low4 - High4	5.49%	1.16

Notes: This table reports the average annualized nearby returns and Newey-West corrected t-statistics for long-short portfolios. Each month, four commodities with the highest signal values are bought (High4) and four with the lowest are sold (Low4). For Media Coverage and Basis, the strategy is reversed (Low4-High4). The results for Emotion Intensity are both statistically and economically significant.

When and Why Does Emotion Intensity Matter?

Cross-Sectional Predictability

To confirm the factor's robustness beyond portfolio sorts, this study uses Fama-MacBeth cross-sectional regressions. The cross-sectional analysis demonstrates that the media emotion intensity factor commands a positive and statistically significant premium, robust to controls for other media-related factors (coverage and sentiment) and established commodity benchmarks. This holds even after including other media factors (coverage and sentiment), and the key commodity benchmarks in the model. This confirms that emotion intensity has cross-sectional predictive power at both the portfolio and individual asset levels.

The Moderating Role of Investor Attention

The premium generated by media emotion intensity is not uniform. It is concentrated in specific, economically intuitive environments. A double-sorting analysis reveals that investor attention, proxied by media coverage levels, significantly moderates the effect of media emotion intensity on returns.

As shown in Table 2, the effect is most powerful for commodities that receive little media attention (low coverage). In this group, the emotion intensity strategy generates a remarkable 12.70% annualized return. Conversely, for commodities already under the media spotlight (high coverage), the strategy's return is statistically close to zero. This is consistent with the limited attention theory. When information is scarce, the emotional framing of the available news has a significant impact on investor judgment. When information is abundant, the effect of emotion is diluted. Similarly, this study finds that the premium is significantly stronger for commodities in backwardation and for those with high return momentum.

Table 2
The Moderating Effect of Media Coverage on the Emotion Intensity Premium

	Single Sort (High - Low)	Double Sort: Conditional on Media Coverage	
		Low Media Coverage	High Media Coverage
Emotion Intensity Premium	6.04%	**12.70%**	**0.01%**
(t-statistic)	*(2.10)*	*(3.39)*	*(0.00)*

Notes: This table reports the annualized return premium of a portfolio that is long on high emotion intensity commodities and short on low emotion intensity commodities. The "Double Sort" columns show the premium is statistically significant only for commodities with low media coverage.

Predicting Trader Behavior and Future Sentiment

Our research identifies a key mechanism linking media emotion intensity to commodity market outcomes, highlighting predictive power for trader behavior and subsequent market sentiment.

- Predicting Trader Behavior: Using Commitment of Traders (CFTC) data, the study finds that media emotion intensity significantly predicts the trading behavior of market participants in the *next* month. Specifically, it precedes an increase in net long positions by non-commercial traders (speculators) and an increase in net short positions by commercial traders (hedgers).

- Predicting Future Sentiment: The study also shows that media emotion intensity in month *t* is a significant predictor of the news sentiment in month *t*+1. Since contemporaneous news sentiment is known to affect returns, emotion intensity acts as a leading indicator.

Validation

To test whether media emotion intensity is a consistent risk premium, this research conducts LASSO (Least Absolute Shrinkage and Selection Operator) regression, a machine learning technique designed to select the most important predictive variables from a large set (following Feng *et al*., 2020). The study tests the emotion intensity factor against all other benchmarks across 200 randomized sub-samples. Across three different LASSO models, the emotion intensity factor is the most frequently selected predictor, appearing in the set of relevant factors up to 100% of the time in the standard LASSO technique. This provides strong evidence that media emotion intensity is a primary driver of returns, not a redundant signal.

Conclusion

This paper introduces media emotion intensity, the ratio of emotional to factual content in news, as a novel and significant predictor of commodity futures returns. It is a distinct factor from news sentiment and media coverage with significant and robust explanatory power.

The key findings are as follows:

1. A long-short portfolio sorted on media emotion intensity generates an economically and statistically significant premium of over 13% per year, outperforming established benchmarks.

2. This premium is robust to transaction costs and is not explained by existing factors, as confirmed by spanning tests, Fama-MacBeth regressions, and its top ranking in LASSO factor selection models.

3. The effect is strongest for commodities with low media coverage, suggesting it is an important signal when investor attention is limited.

4. The media emotion intensity signal forecasts future news sentiment and the resulting positioning of both speculative and commercial traders.

For investors, incorporating the media emotion intensity factor into asset allocation and risk management models can significantly improve predictive accuracy and enhance portfolio performance. For policymakers, understanding the influence of emotional media content on market behavior is important for maintaining stability. The study suggests that in commodity trading, it is not just what is reported that matters, but how it is reported.

References

Beckett, C., 2015, "How Journalism is Turning Emotional and What that Might Mean for News." Accessed via website: https://blogs.lse.ac.uk/polis/2015/09/10/how-journalism-is-turning-emotional-and-what-that-might-mean-for-news/ on December 4, 2023.

Boons, M. and M. P. Prado, 2019, "Basis-Momentum," *Journal of Finance*, Vol. 74, No. 1, February, pp. 239-279.

Fang, L. and J. Peress, 2009, "Media Coverage and the Cross-Section of Stock Returns," *Journal of Finance*, Vol. 64, No. 5, October, pp. 2023-2052.

Feng, G., Giglio, S. and D. Xiu, 2020, "Taming the Factor Zoo: A Test of New Factors," *Journal of Finance*, Vol. 75, No. 3, June, pp. 1327-1370.

Loewenstein, G., 1996, "Out of Control: Visceral Influences on Behavior," *Organizational Behavior and Human Decision Processes*, Vol. 65, No. 3, March, pp. 272-292.

Szymanowska, M., De Roon, F., Nijman, T. and R. Van Den Goorbergh, 2014, "An Anatomy of Commodity Futures Risk Premia," *Journal of Finance*, Vol. 69, No. 1, February, pp. 453-482.

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, No. 3, June, pp. 1139-1168.

Vuilleumier, P., 2005, "How Brains Beware: Neural Mechanisms of Emotional Attention," *Trends in Cognitive Sciences*, Vol. 9, No. 12, December, pp. 585-594.

Keywords

Commodity futures, media emotion intensity, asset return, LASSO.

Author Biographies

YEGUANG CHI, Ph.D.
University of Auckland, New Zealand

Dr. Yeguang Chi is a Senior Lecturer at the Department of Accounting and Finance, University of Auckland. Dr. Chi's research focuses on the areas of asset pricing, information and markets, institutional money management, Chinese financial markets, and cryptocurrency.

LINA EL-JAHEL, Ph.D.
University of Auckland, New Zealand

Dr. Lina El-Jahel is an Associate Professor of Finance at the Department of Accounting and Finance, University of Auckland. Her research focuses on the areas of credit risk, banking, behavioral finance and asset pricing.

THANH VU, Ph.D.
Auckland University of Technology, New Zealand

Dr. Thanh Vu is a Research Fellow at the Auckland Centre for Financial Research (ACFR), Auckland University of Technology, where his research focuses on asset pricing, behavioral finance, corporate finance, and the integration of machine learning and AI in financial research.

Forecasting Realized Volatility of Crude Oil Futures Prices Based on Machine Learning

Jiawen Luo, Ph.D.
School of Business Administration, South China University of Technology, China

Tony Klein, Ph.D.
Faculty of Business and Economics, Technische Universität Chemnitz, Germany

Thomas Walther, Ph.D.
Utrecht School of Economics, Utrecht University, The Netherlands; and Editorial Board Member, *Commodity Insights Digest*

Qiang Ji, Ph.D.
Institutes of Science and Development, Chinese Academy of Sciences, China

Published in *Journal of Forecasting* and Available at: https://onlinelibrary.wiley.com/doi/10.1002/for.3077

Extending the popular heterogeneous autoregressive (HAR) model with additional information channels to forecast realized volatility of WTI futures prices, we show that machine learning-generated forecasts provide better forecasting quality and that portfolios that are constructed with these forecasts outperform their competing models resulting in economic gains. Analyzing the selection process, we show that information channels vary across forecasting horizons. Variable selection produces clusters and provides evidence that there are structural changes with regard to the significance of information channels.

Introduction

Modelling and forecasting volatility is a crucial task in risk management to quantify *e.g.,* commodity market risk exposure. One of the tools to predict asset price fluctuations is the heterogeneous autoregressive (HAR) model (Corsi, 2009). Structurally a simple, linear model of three different horizons of realized volatility, it has been proven hard to beat by more complex models in out-of-sample settings (Degiannakis *et al*., 2022). In machine learning (ML) terminology, the more complex models typically rely on features based on returns, *e.g.,* asymmetry between positive and negative returns or jump processes. In this study, we extend the baseline HAR model with exogenous information. In particular, we include equity, sentiment, other commodities, attention, and foreign exchange information channels.

In addition, we use ML techniques to improve the HAR model and better select relevant predictors. In this way, the inclusion of certain information channels is purely data-driven and not at the discretion of the forecaster. We show that the ML variants can outperform the baseline model both in accuracy of the predictions and economic gains. Building on the time-varying selection of predictors, we further analyze the importance of the information from different channels. Particularly, foreign exchange information adds value beyond realized volatility.

The main contribution of this paper is to provide new evidence on the role of diverse exogenous information channels, including equity, commodity, FX, sentiment, and Google Trends, in explaining the realized volatility of WTI crude oil futures. Extending the baseline HAR model with 84 predictors, we apply

and compare multiple ML-based variable selection methods (LASSO, BMA, BAG, and SSVS) to address parameter proliferation and multicollinearity. We show that incorporating selected exogenous variables improves both forecasting accuracy and economic value, with FX markets emerging as particularly important. Our results further reveal that the number of relevant channels increases with the forecasting horizon, contributing to the literature on market connectedness and HAR model forecasting.

Methodology and Data

We model realized volatility using high-frequency intraday data and extend the heterogeneous autoregressive (HAR) framework to incorporate a broad range of exogenous information channels. Following Corsi (2009), we compute daily realized volatility (RV) as the sum of squared intraday returns sampled at 5-minute intervals:

$$RV_t = \sum_{i=1}^{m} r_{t,i}^2$$

where $r_{t,i} = 100 \cdot (\log(P_{t,i}) - \log P_{t,i-1})$. Our baseline model is the standard HAR specification, in which the forecast of log realized volatility over a given horizon *h* depends on daily, weekly (5-day), and monthly (22-day) averages of past log realized volatility. We estimate separate regressions for each forecasting horizon rather than imposing fixed weights across horizons, allowing for greater flexibility in capturing horizon-specific dynamics:

$$\log RV_{t+h}^{(h)} = \beta_0 + \beta_1 \log RV_t + \beta_2 \log RV_t^{(5)} + \beta_3 \log RV_t^{(22)} + u_{t+h}^{(h)},$$

with

$$\log RV_{t+h}^{(h)} = \frac{1}{h}\sum_{j=1}^{h} \log RV_{t+h-j+1}.$$

In this paper, we forecast the volatility of the WTI crude oil futures. For forecasts beyond 1-step ahead, our models predict the average realized volatility over the forecasting horizon, which we use as a proxy for the realized volatility h-steps ahead. To enrich the information set, we extend the baseline HAR by including variables from five exogenous predictor categories: (i) equity market volatilities from major global indices: Dow Jones Industrial Average, CAC40, FTSE, DAX30, HSI, NIKKEI225, S&P 500, SSE, and EUROSTOXX; (ii) sentiment and uncertainty indicators: the Global Financial Stress Index, VIX, Oil Volatility Index OVX, and Economic Policy Uncertainty; (iii) commodity futures volatilities, including Natural Gas, Corn, Soybeans, and Gold; (iv) foreign exchange market volatilities: U.S. weighted currency basket, EUR/USD, Yen/USD, Yuan/USD, Ruble/USD, and GBP/USD; and (v) Google Trends search volumes to gauge the attention to “Oil Production”, “Financial Crisis”, “Oil Demand”, “Oil Price”, and “OPEC Conference”. Given the five predictor sets, we construct HAR-Equity, HAR-Sentiment, HAR-Commodity, HAR-FX and HAR-Google models respectively. For Equity, we use realized volatilities sampled at a matching 5-min interval. For the commodity futures and FX, we proxy volatilities by using squared daily returns. For each variable, we construct daily, weekly, and monthly averages to match the HAR structure, which results in a

total of 84 exogenous predictors when all categories are combined. We estimate category-specific HAR extensions as well as two "kitchen-sink" approaches: a comprehensive "HAR-All" specification that simultaneously incorporates all predictors; and a specification which first runs a Principal Component Analysis and only uses the first principal component of each information channel and each horizon and we define it as HAR-factor model. Lastly, we also add an Autoregressive (AR) model of order 1 to the set of models to have a very simple benchmark.

The large number of predictors in these extended models raises issues of over-parameterization and multicollinearity, both within and across information channels. To address this, we employ several ML–based techniques that are well-suited to high-dimensional forecasting problems. Specifically, we use the Least Absolute Shrinkage and Selection Operator (LASSO), Bayesian Model Averaging (BMA), Bagging, and Stochastic Search Variable Selection (SSVS). While LASSO and BAG reduce dimensionality, the BMA and SSVS use a weighting of predictors to emphasize the most important variables.

Results

We use the above models to forecast oil futures volatility (WTI) for the next day, the next week (5 days), and the next month (22 days). In the first part of the results, we analyze the predictive performance in terms of accuracy, *i.e.,* we compare the mean squared and mean absolute errors of the different model configurations. In order to understand which models actually perform better than others (in statistical terms), we use the Model Confidence Set (MCS, Hansen *et al.*, 2011). Given a confidence level, the MCS is a subset of models with equal predictive performance.

In the first set of our results, we find that most model-predictor combinations have equal predictive performance when it comes to the mean squared error. Only for mean absolute errors, we find a more exclusive set if we lower our confidence (we use confidence sets of 10% and 25%, with 25% being more exclusive but less confident). Especially for the 1-day ahead forecast, the MCS only contains four models from the set of 42 models: BMA- and BAG-HAR-Google, LASSO- and BAG-HAR-All.

For the 1-week ahead forecast (5 days), we find mostly LASSO models to be included in the 25% MCS of both MSE and MAE. While the LASSO-HAR-FX combination has the lowest predictive errors, its performance is not statistically different from the LASSO-HAR, *i.e.*, the additional information from the FX channel only plays a marginal role.

For the 22-day ahead (1-month) forecasts, statistical accuracy is very similar across models. For the MSE, even the more restrictive 25% confidence set contains all HAR variants. The MAE 25% set mainly includes the LASSO-based models together with a few others. For longer horizons, the Google-augmented variants appear to be the best-performing models.

After looking at the statistical accuracy, we are interested in the economic value of the forecasts. Particularly, we compare the values to the standard AR with and without information from either Google attention or other asset classes. In this way, we can assess both the addition of using ML techniques and the addition of extra information over a very naïve benchmark. We assume an investor with a quadratic

utility function with either mild ($\gamma = 1$) or very strong ($\gamma = 10$) risk aversion and compare the models over the forecasting horizons of one month. For each model-predictor combination, we forecast the next step's volatility. In order to predict the next period's return, we use a simple 1-year moving average for all combinations. Then the allocation to the oil futures (the risky asset) is $\widehat{\omega}_t = 1\backslash\gamma \left(\frac{\widehat{r_{t+1}}}{\widehat{\sigma}^2_{t+1}}\right)$. The portfolio return comprises the realized return from the investment in the oil futures plus the investment of the remainder into the 5-year U.S. Treasury Note.

Our findings show that while ML-augmented HAR models outperform standard HAR models in terms of statistical accuracy, this advantage does not always carry over to economic value. In portfolio applications, the BMA-based models achieve the greatest economic gains. The BMA-HAR-FX yields the highest annualized return and economic value compared to both AR(1) and standard HAR benchmarks. Risk-adjusted performance, however, is strongest for BAG models, with the BAG-HAR-All achieving the highest Sharpe ratio. LASSO models, on the other hand, do not translate statistical accuracy into higher utility.

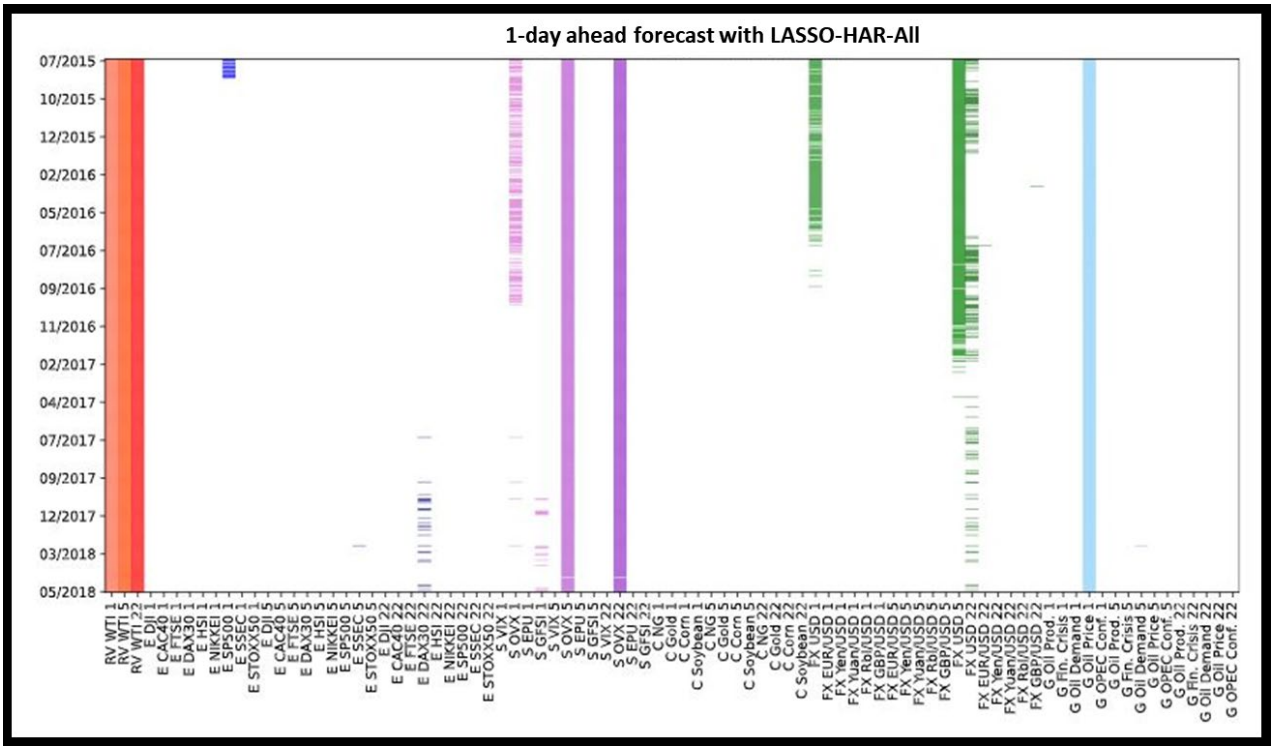
We observe two general trends in our results. Using additional information channels (especially from sentiment, commodity markets, or FX) lead to higher economic gains. ML-models can then further improve the forecasts, but only marginally.

Our last set of results looks at the time-varying drivers of the forecasts. For this exercise, we use the LASSO-HAR-All specification which contains every information channel at three different horizons. This allows us at every forecasting step to assess which of the predictors the LASSO selected over our out-of-sample period from 2015 to 2018. In this way, we can answer which are the most important variables, how sticky they are over time, and if additional information channels are able to replace the information from the previous realized volatility in the HAR model. See the three plots in Figure 1 below.

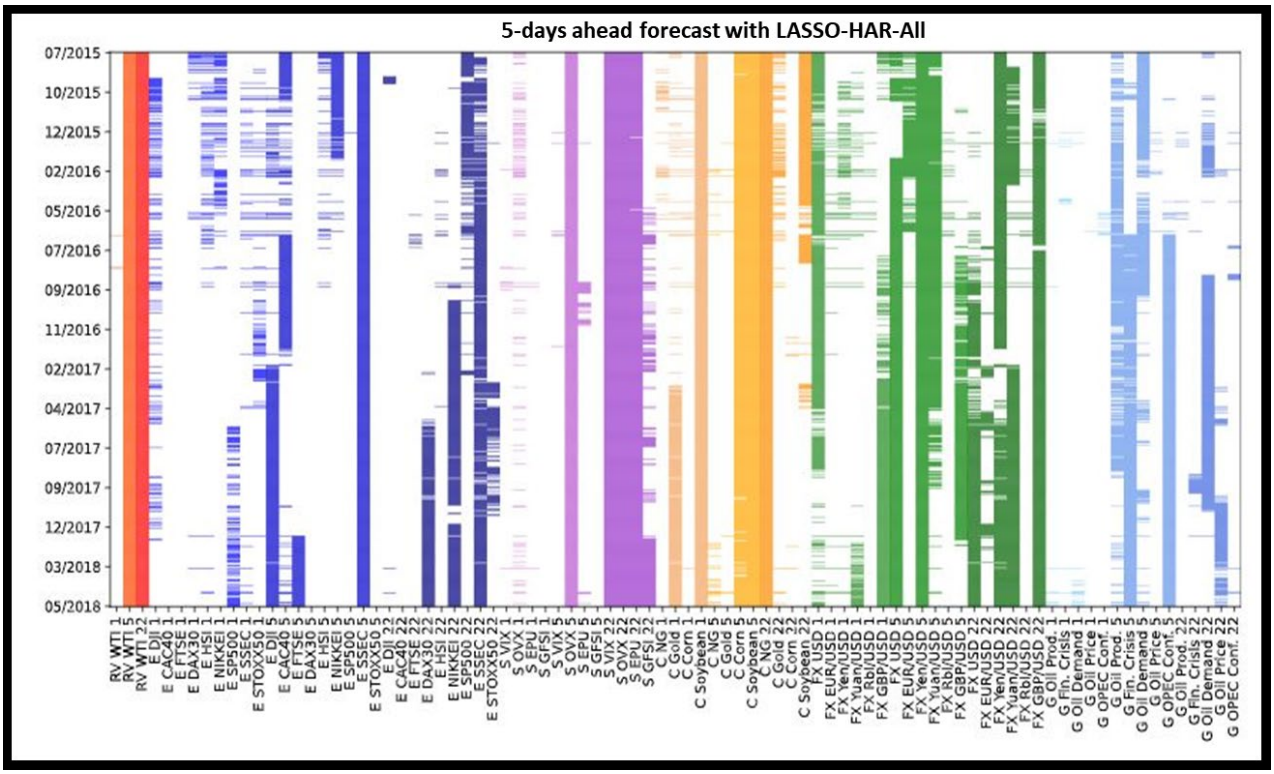
Three main observations emerge: (1) the information from the previous volatility remains important and is not replaced by any drivers. Only in the 5-day ahead model, we find that the last day's realized volatility is not important for the model. (2) While only a few predictors are selected for the 1-day ahead forecasts, we find that the longer the forecasting horizons, the more predictors are selected by the LASSO. (3) Apart from a few very persistent variables over time (*e.g.,* OVX and attention to Oil Price for 1-day; SSEC, OVX for 5-days; and NIKKEI and OVX for 22-days), the drivers vary a lot over time. We even observe clustering of some variables, meaning they are persistent for a short period of time before falling out of the selection again. For example, in the 22-days forecast, short-term information from the stock markets were selected until 2016 and fell out almost entirely afterwards.

Figure 1

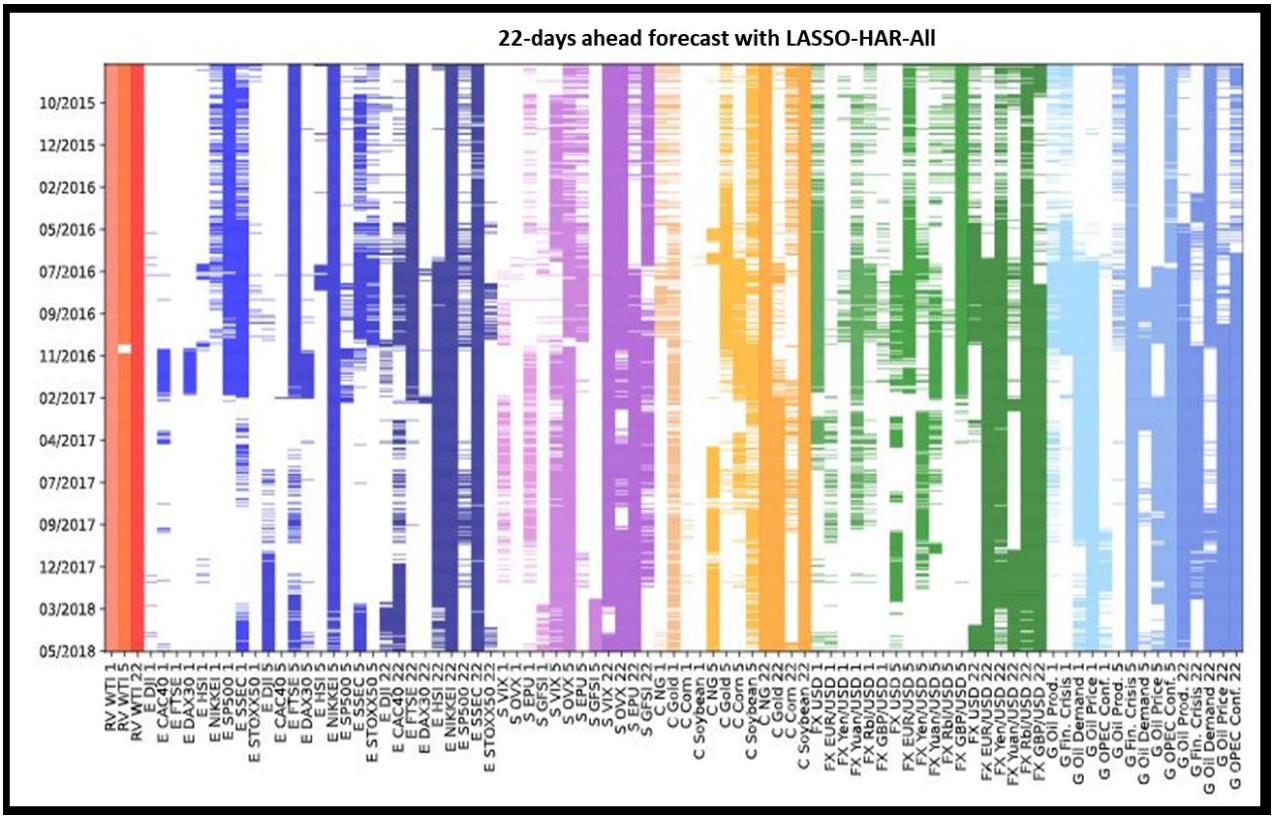
(a) Variable Selection Over Time with the LASSO-HAR-All (1-day ahead forecast)



(b) Variable Selection Over Time with the LASSO-HAR-All (5-days ahead forecast)



(c) Variable Selection Over Time with the LASSO-HAR-All (22-days ahead forecast)



Lastly, we also compare the variable importance. In short, the variable importance is the absolute, standardized coefficient for a predictor over the sum of all absolute, standardized coefficients from a LASSO regression. We then take the average for each predictor over the entire out-of-sample period. See Figure 2, which is further discussed on the next page.

Figure 2

Ranked Variable Importance of Individual Variables (Top 10) Across the Applied Forecasting Horizons

Rank	1 day	VI	5 days	VI	22 days	VI
1	RV WTI 22	47.8%	RV WTI 22	53.6%	RV WTI 22	66.4%
2	RV WTI 5	30.9%	FX Yuan/USD 22	9.4%	FX EUR/USD 22	5.6%
3	RV WTI 1	8.9%	RV WTI 5	7.4%	FX Yuan/USD 22	4.1%
4	FX USD 5	5.1%	S OVX 22	5.2%	FX Yen/USD 22	2.8%
5	S OVX 22	3.5%	FX Yen/USD 5	4.6%	E NIKKEI 22	2.7%
6	S VIX 5	2.0%	FX Yuan/USD 5	4.3%	E HSI 22	2.6%
7	FX USD 22	0.7%	FX GBP/USD 22	3.0%	S GFSI 22	1.6%
8	FX USD 1	0.6%	FX USD 5	2.7%	RV WTI 5	1.5%
9	G OiL Price 1	0.5%	FX USD 22	2.2%	FX GBP/USD 22	1.3%
10	S OVX 1	<0.1%	FX Yen/USD 22	2.0%	C Soybean 22	1.3%

We first look at all individual predictor-horizon combinations and find the 22-day realized volatility to be the most important variable over all three forecasting horizons. It makes up 47.8% in the 1-day ahead case up to 66.4% in the 22-day ahead forecast. For the 1-day ahead prediction, the order is 22-day RV, 5-day RV, and 1-day RV. For the longer forecasting horizons, the previous day's RV falls entirely out of the top 10 predictors. Predictors from other channels contribute more to the forecasts. Within the top 10 predictors, we find FX and sentiment predictors over all forecasting horizons. Other information channels appear also, but only with incremental contribution (*e.g.,* Google attention to Oil Price for 1-day ahead, equity market for 22-days ahead).

Finally, we condense the information channels regardless of their horizon and look at their importance over the entire period. As stated above, the past realized volatility is by far the strongest contributor to the prediction. Interestingly, FX is the runner-up for all horizons with contributions between 6.3% (1-day) to even 28.8% (5-days). Information from other commodities and Google attention play a minor role for the oil price volatility forecast. In the 1-day ahead case, *e.g.*, they are never selected.

Conclusions

This paper shows that extending HAR models with exogenous information and recent ML techniques improves both forecasting accuracy and portfolio performance. ML-based variable selection, particularly through LASSO, improves statistical accuracy, while BMA models provide the greatest economic gains. We find that the relevance of predictors is horizon-dependent: short-term forecasts are dominated by past volatility and use only a few additional information channels. For longer horizons, sentiment, equity, commodity, and especially FX markets have more importance without replacing realized volatility. Google-based indicators play a role but remain of limited importance compared to traditional financial channels. Overall, our results highlight that while past realized volatility remains the most important predictor, incorporating broader information channels and ML-driven selection yields meaningful improvements in both predictive and economic value.

References

Corsi, F., 2009, "A Simple Approximate Long-Memory Model of Realized Volatility," *Journal of Financial Econometrics*, Vol. 7, No. 2, Spring, pp. 174-196. https://doi.org/10.1093/jjfinec/nbp001.

Degiannakis, S., Filis, G., Klein, T. and T. Walther, 2022, "Forecasting Realized Volatility of Agricultural Commodities," *International Journal of Forecasting*, Vol. 38, No. 1, January-March pp. 74-96. https://doi.org/10.1016/j.ijforecast.2019.08.011.

Hansen, P. R., Lunde, A. and J. M. Nason, 2011, "The Model Confidence Set," *Econometrica*, Vol. 79, No. 2, March, pp. 453-497. https://doi.org/10.3982/ECTA5771.

Keywords

Forecasting, crude oil, realized volatility, exogenous predictors, machine learning.

Author Biographies

JIAWEN LUO, Ph.D.
School of Business Administration, South China University of Technology, China

Jiawen Luo is an Associate Professor at the South China University of Technology. Her research interests include financial economics, financial risk analysis, and energy finance. Her research is published in journals such as *International Journal of Forecasting*, *Journal of Forecasting*, *Journal of Futures Markets*, *Energy Economics*, and *European Journal of Finance*.

TONY KLEIN, Ph.D.
Faculty of Business and Economics, Technische Universität Chemnitz, Germany

Tony Klein is a Professor of Finance and holds the Chair of Finance and Banking at Technische Universität Chemnitz, Germany. He previously held positions at Queen's University Belfast, U.K., and Universitat de Barcelona, Spain. His research interests include econometric modelling and machine learning within the fields of finance and economics, energy and carbon economics, and the financialization of commodity markets and connectedness to financial markets.

THOMAS WALTHER, Ph.D.
Utrecht School of Economics, Utrecht University, The Netherlands and Technische Universität Dresden, Germany

Thomas Walther is an Associate Professor of Finance with the Utrecht School of Economics, Utrecht University, The Netherlands. Formerly, he was an Assistant Professor for Energy Finance (Non-Tenure Track) at the School of Finance, University of St. Gallen, Switzerland. He is also a Research Fellow for Finance at Technische Universität Dresden, Germany from which he received a B.Sc. (2010) and M.Sc. (2013) in Industrial Engineering and Management and a Ph.D. (2017) in Business and Economics. His research focuses on commodity/energy/climate finance. He initiated the European Sustainable Finance Ph.D. Workshop in 2023 in Utrecht (with more information about the latest workshop available here: http://www.esfpw.eu). He is also the co-founder of the Dutch Sustainable Finance Network (http://www.dsfn.eu), which has organized an annual workshop since 2024.

QIANG JI, Ph.D.
Institutes of Science and Development, Chinese Academy of Sciences, China

Qiang Ji is the executive director of the support center for discipline research in the Institutes of Science and Development, Chinese Academy of Sciences (ISDCAS). He presides over a program supported by the Excellent Young Scientists Fund of the National Natural Science Foundation of China and is granted the Excellent Member by the Youth Innovation Promotion Association of CAS. He has been engaged in research work in the fields of energy strategic management, energy and climate finance, energy forecasting, and risk management. He is the co-founder of the China Energy Finance Network and the vice president of the International Society for Energy Transition Studies and Society for the Studies of Climate Finance. He is a founding editor of the *Journal of Climate Finance* and serves as senior editor and associate editor for over 10 energy and financial journals, such as *Energy Economics*, *Finance Research Letters*, and *International Review of Financial Analysis*. He has published more than 150 SCI/SSCI indexed articles in peer-reviewed journals and has been named in the global "Highly Cited Researchers" lists in the year 2021–2023 released by Clarivate Analytics and "World's Top 2% Scientists" lists released by Stanford University in the year 2020–2023.

From Theory to Practice: Myths and Mysteries About Speculation in the Oil Market

Ilia Bouchouev, Ph.D.
Managing Partner, Pentathlon Investments LLC, U.S.A. and Editorial Board Member, *Commodity Insights Digest*

Wu-Yen Sun
Visiting Researcher, King Abdullah Petroleum Studies and Research Center (KAPSARC), KSA

Available at: https://www.kapsarc.org/our-offerings/publications/myths-and-mysteries-about-speculation-in-the-oil-market/

The article provides unique quantitative insights about highly secretive and poorly understood speculation in the oil market. To demystify the behavior of speculators, we look at the problem from several different angles. First, we explain how the presence of the large over-the-counter (OTC) oil derivatives market leads to mischaracterization of traditional hedgers and speculators. We then explain what makes oil speculation special and different from speculation across many other commodities. Consequently, we identify the winners and the losers in the oil futures market. To model the behavior of the winners which we associate with fast-moving quantitative hedge funds, we develop a novel framework based on a simple neural-network algorithm. We conclude by analyzing a popular investment strategy of following the winners, or the hedge-funds' flows.

Introduction

Ever since the oil market was born, speculation has been an integral part of it. This fact is undeniable, regardless of our wishes, beliefs, or the limitations of conventional empirical methods used to detect it. So why is speculation so difficult to observe? One challenge is the limited access that academic researchers have to timely and relevant data, coupled with an incomplete understanding of the motivations driving many market participants. Another is the reluctance of professional traders to share insights, given the proprietary nature of their trading strategies. This article aims to bridge that gap.

The conventional framework of modelling speculation in commodity markets relies on futures positioning data provided by regulators in the form of the so-called Commitments of Traders (COT) reports. Various versions of COT reports have been in existence for agricultural commodities since the 1920s. Their primary objective is to differentiate between futures and options positions held by hedgers and speculators, which were labelled as commercial (COM) and non-commercial (NC) traders. The presumption is that the COM traders deal with the physical commodity and represent hedgers. In contrast, the NC traders do not handle physical deliveries, thus they are deemed to be speculators.

While this simple categorization may be adequate for certain commodities, it has proven to be of limited value for analyzing participants in a more complex ecosystem of the oil market where speculation and hedging are more intertwined. For example, sophisticated commercial traders that own or lease physical assets, such as storage, pipelines or oil tankers, are also known to be large speculators, as they aim to leverage an informational edge obtained from their ownership of assets. Vice versa, some non-commercial financial investors, which are typically labelled as speculators, often hold oil futures as a hedge against inflation and geopolitical risks for their broader financial portfolios. Finally, in contrast to many other commodities, oil has a large over-the-counter (OTC) market where oil producers and consumers hedge

their price exposures, whereas futures are effectively held on their behalf by dealers classified as commercial traders.

Ilia Bouchouev, Ph.D. (front row center in the light blue shirt), lectured at Bayes Business School (U.K.) on June 12, 2024. Dr. Bouchouev's workshop covered "Quantitative Trading in the Oil Market." He is the Managing Partner of Pentathlon Investments LLC (U.S.A.) and an Editorial Board member of the *Commodity Insights Digest*.

In an attempt to provide more transparency, the U.S. Commodity Futures Trading Commission (CFTC) revamped the format of its positioning reports. Starting in 2006, the COM category was split into two new subgroups: producers, merchants, processors, and users (PMPU); and swap dealers (SD). The NC category was also split into managed money (MM) and other (OTH) category. These reports became known as the Disaggregated Commitments of Traders (DCOT) reports. While this decomposition brought clarity for many commodities that lack large OTC markets, DCOT reports have not resolved the problem of participant identification in the oil market. In fact, they may have introduced even more confusion by effectively mislabeling some of the largest traders in the market.

The explicit reference to producers under the PMPU category only reinforced the incorrect association between genuine oil producers and the perception of their hedging activity. Since actual hedging by producers occurs predominantly OTC via bilateral agreements with their banks, any evidence of producer hedges typically appears in the SD category. In contrast, the PMPU label in the oil market has become something of a misnomer, as this category is, in reality, still dominated by highly sophisticated physical oil traders who have acquired physical assets primarily to enhance their speculative activity.

In theory, one can deduce some useful information about producer hedging by analyzing the SD category instead, as swap dealers trade with producers OTC and then offset their risks in the futures market. However, swap dealers manage fairly diverse portfolios. They also hold long futures positions as hedges

against OTC agreements with consumer-type clients, such as airlines and utilities. More importantly, they also hold long futures positions against OTC commodity index swaps with pension funds, insurance companies and other financial investors. Long futures positions held on behalf of investors by dealers tend to offset their short positions held on behalf of producers. This makes the SD category a "catch-all" bucket with rather limited informational value.

The DCOT reports did provide some helpful granularity for positions held by non-commercial traders, often associated with financial speculators. To characterize this non-homogeneous group, it is helpful to split it into three types of speculative activity.

The first group consists of large institutional investors who hold oil derivatives as part of diversified investment portfolios. Their positions are predominantly long, as they seek to hedge against inflation risk. While some investors hold futures directly, others prefer to gain exposure to commodity prices through OTC swaps with banks – similar to how oil producers hedge. As a result, the positions held by these so-called passive investors may appear either in the MM category if they trade futures directly or in the SD category if they transact OTC and futures are reported by dealers. The exact split between these two channels is difficult to estimate due to minimal reporting requirements and the highly confidential nature of bilateral OTC transactions.

The second – and likely the most influential – group of non-commercial speculators in today's oil market comprises quantitative hedge funds, or commodity trading advisors (CTAs). These funds rely heavily on algorithmic trading, executing strategies based on quantitative models with little or no human intervention. Quantitative funds have seen tremendous growth over the past decade with the proliferation of new data sources and the continued digitalization of financial markets. Unlike passive investors, CTAs adjust their oil positions rapidly, often contributing to significant short-term price fluctuations. Arguably, this group has become the largest driver of short-term oil prices, and shedding light on their behavior is one of the primary objectives of this paper.

The third category of speculators consists mostly of discretionary traders, including many retail investors. Fortunately, the DCOT reports – while not particularly helpful in clarifying the roles of commercial traders and hedgers – have proven more useful in understanding speculators. Since the nature of oil speculators reported under the NC category is highly heterogeneous, the added granularity provided by the DCOT reports has offered some valuable new insights which we build upon in the following section.

Who are the Winners and Losers in the Oil Futures Market?

Many academic studies of the behavior of participants in commodity markets revolve around the Keynes-Hicks theory of "normal backwardation."[1] This theory asserts that commodity futures markets are naturally in disequilibrium as hedging needs of commodity producers structurally exceeds those of commodity consumers. Thus, futures prices must trade at a discount to the expected spot price. This discount then provides incentives to speculators to take the other side of the imbalanced futures market. Speculators are then expected to realize gains over time by buying and holding commodity futures.

Given the very convoluted nature of hedging and speculation in the oil market, it should not come as a surprise that this theory does not hold well in the oil market. To illustrate, we take actual weekly positions held by trader groups and calculate their annualized profit and loss (P&L) and the annualized standard deviation (STDEV) of P&L directly in dollar terms. We then use the ratio of the two to measure risk-adjusted performance, similar to the Sharpe ratio (SR).[2] In contrast to many academic studies that use somewhat ambiguous measures of futures returns and normalized measures of speculative and hedging pressure in the context of broader commodity portfolios, our direct P&L-based approach highlights the magnitude of gains and losses and uses metrics commonly employed by traders in the oil market.

We compute these realized performance metrics for the two legacy categories – NC and COM traders in COT reports – as well as for the four participant categories in the DCOT reports. For benchmarking purposes, we include the results of a static rolling long position of 100,000 contracts (equivalent to 100 million barrels).

We use WTI data from June 2006 to June 2024 and Brent data from January 2011 to June 2024, corresponding to the availability of DCOT reports published by the CFTC and ICE, respectively. To enable a direct comparison of the behavior of traders across both markets over the same period, we also present results for WTI for a shorter sample aligned with the Brent DCOT data availability. We assume all trades are rolled to the next futures maturity at the end of each calendar month. Table 1 illustrates our results.

Table 1
Profitability of Different Groups of Traders

	COM	NC	PMPU	SD	MM	OTH	MM-SLOW	MM-FAST	STATIC LONG
a) WTI (2006-2024)									
Annualized P&L	112,747,800	(148,577,800)	266,499,600	(153,751,800)	341,154,900	(489,732,700)	(178,515,000)	519,669,900	(165,550,800)
STDEV	7,883,983,000	7,312,810,000	3,192,569,000	7,871,738,000	4,760,920,000	3,044,494,000	4,847,470,000	1,103,688,000	2,577,695,000
Sharpe Ratio	0.01	(0.02)	0.08	(0.02)	0.07	**(0.16)**	(0.04)	**0.47**	(0.06)
b) BRT (2011-2024)									
Annualized P&L	(42,994,600)	(42,206,220)	(437,904,500)	394,909,900	881,772,700	(923,978,900)	106,674,900	775,097,800	39,088,280
STDEV	2,972,062,000	2,773,672,000	9,253,212,000	7,117,329,000	5,021,280,000	3,269,173,000	5,387,889,000	1,444,252,000	2,386,923,000
Sharpe Ratio	(0.01)	(0.02)	(0.05)	0.06	0.18	**(0.28)**	0.02	**0.54**	0.02
c) WTI (2011-2024)									
Annualized P&L	487,986,700	(425,292,900)	616,907,000	(128,920,400)	64,177,990	(489,470,900)	(630,074,000)	694,252,000	(264,411,800)
STDEV	8,986,299,000	8,350,252,000	2,617,513,000	8,932,683,000	5,258,862,000	3,482,721,000	5,396,155,000	1,162,292,000	2,430,607,000
Sharpe Ratio	0.05	(0.05)	0.24	(0.01)	0.01	**(0.14)**	(0.12)	**0.60**	(0.11)

Since futures trading is a zero-sum game, does this mean that no participants consistently make or lose money and that there are no structural biases driven by the behavior of specific market participants? This is not necessarily the case. To answer this, we need to dig a little deeper to untangle the complex puzzle of trader behavior. In order to derive more meaningful conclusions about the behavior of various market participants – particularly speculators – one approach is to segment them based on trading frequency.

The idea of classifying traders according to how frequently they trade and how long they hold positions is not new. For example, Kang *et al.* (2020) applied this approach to study the elusive Keynesian risk premium across a diversified portfolio of commodities. They concluded that hedgers and speculators

provide liquidity to each other on different time scales: in the short term, speculators consume liquidity and pay a risk premium to hedgers; in the long term, speculators tend to collect the risk premium paid by hedgers. We applied the results of their study specifically to the oil market. While we found some evidence of a short-term liquidity risk premium paid by speculators during the earlier period of oil futures trading, this phenomenon appears to have diminished over time. We also confirmed that the long-term hedging premium in oil is statistically insignificant.

Our approach is instead to decompose the crucial MM (managed money) category into MM-SLOW and MM-FAST. MM-SLOW is defined as the three-month moving average of MM positions, while MM-FAST represents the residual – capturing faster trading positions established relative to the three-month average baseline. The decomposition of the MM category into fast and slow components is essential. It allows us to use MM-SLOW to proxy the behavior of passive investors and MM-FAST to represent more dynamic trading by quantitative hedge funds. Notably, the average of MM-FAST positions is approximately zero, reflecting the typical long-short trading patterns of quantitative hedge funds, which are just as likely to be long as short. This also removes the influence of any price bias in the sample.

One can see that neither COM nor NC traders consistently generate profits. This confirms our earlier observation about the absence of any structural hedging or speculative risk premia in the oil market. Furthermore, even when the data are disaggregated into the four DCOT categories, none of the groups consistently produce attractive profits. Their individual Sharpe ratios are economically insignificant.

However, the further segmentation of MM traders into MM-SLOW and MM-FAST reveals a more compelling insight. The performance of MM-FAST – our proxy for the trading behavior of quantitative hedge funds – is notably profitable. On average, MM-FAST traders generate relatively high Sharpe ratios between 0.47 and 0.60. This highlights that profitable trading opportunities in the oil market tend to be short-lived and cannot be captured by passive investment strategies seeking structural risk premia. For a more detailed discussion of the behavior of other traders, we refer to the full text of the article.

What is clear from the results in Table 1 is that quantitative funds – represented by MM-FAST – possess unique skills that other participants do not. They are able to apply these skills consistently to generate highly significant risk-adjusted returns in both the WTI and Brent markets, regardless of the price bias in the sample. In the next section, we explore what these elusive skills might be – and whether a more casual oil trader can, in some way, replicate their behavior.

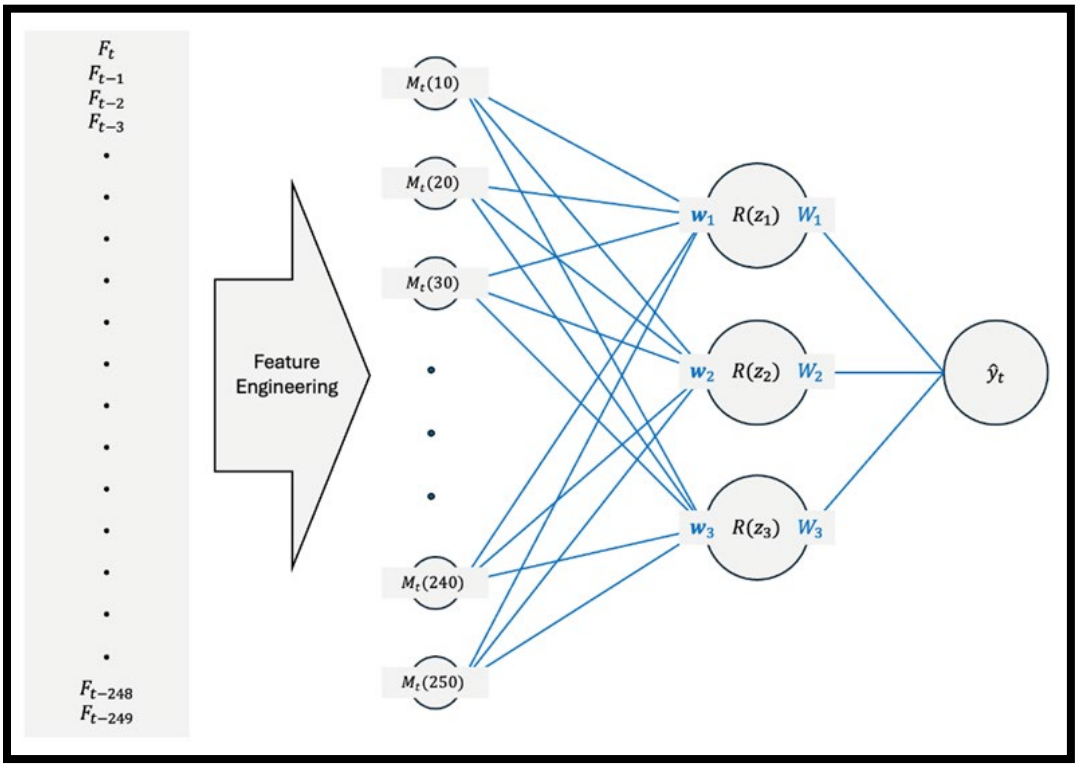
How Can One Replicate the Strategy of the Winners?

In this section, we show how the behavior of quantitative hedge funds can be effectively replicated. It is well known that many quantitative funds favor various momentum strategies, but the exact details of these strategies – especially the dynamic sizing of their positions – remain highly confidential. Obviously, we cannot reconstruct individual hedge funds' trading frameworks, but we demonstrate that valuable insights can be gained about their collective behavior.

This problem is particularly well suited for the application of a mathematical technique known as a neural network. In this framework, historical prices serve as inputs or features, and quantitative funds can be interpreted as neurons, which are activated through a nonlinear function. The process of training the neural network is effectively a form of reverse-engineering the decision-making behavior of hedge funds.

The overarching goal of the neural-network framework is to construct a nonlinear function, $y = f(x)$, that maps the vector of inputs x (known as features) into another vector of observable outputs y (the target). In our case, the features are historical oil prices or some combination of them, and the target is the futures position held by MM-FAST traders at a given time. Since quantitative hedge funds favor momentum-based trading, we use moving averages of prices to construct momentum signals and use those signals as input features – a process known as feature engineering. Figure 1 illustrates our neural network schematically.

Figure 1
The Neural-Network Representation of Momentum Trading



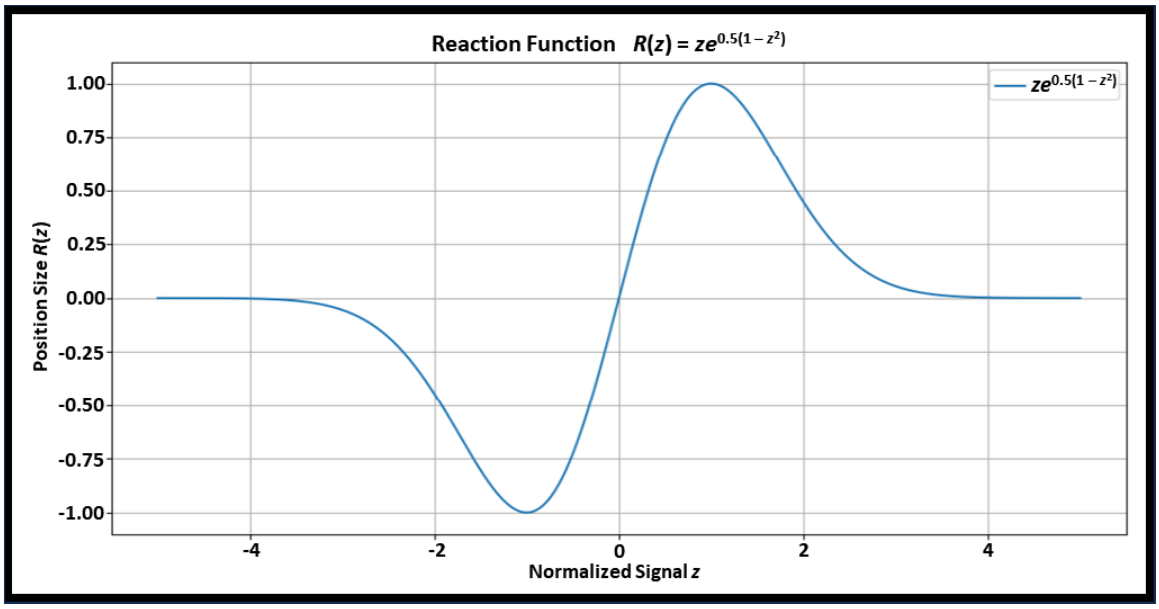
This neural network mimics the behavior of quantitative hedge funds extremely well. In fact, one could even argue that what hedge funds do in practice is a classical definition of a neural network. First, the prices are "re-engineered" into moving averages which are used to construct the momentum signals. Then various linear combinations of basic momentum signals are constructed to form several nodes. The nodes are called neurons, and such construction is designed to mimic the functioning of a human brain. In our diagram, we use three neurons, indexed as $k = 1,2,3$, and the input to each neuron generated by the weighted combination of the basic momentum signals is denoted by z_k. One can think about our network

as a representation of an industry aggregate hedge fund trading a portfolio of momentum strategies on three different frequencies.

The incoming weights for each of the three neurons $w_{k,n}$ depend on 25 momentum signals $M_t(n)$ where $n = 10,20, \dots, 250$ refer to the lookback period.[3] In other words, momentum signals could be understood as a vector of 25 features. These linear combinations of features form the so-called hidden layer. In our example, we only use one hidden layer with three neurons. More complicated networks typically contain more neurons and multiple hidden layers. However, with one hidden layer, the neurons can be interpreted as quantitative hedge funds focusing on momentum signals with three different frequencies, the practice commonly adopted in the industry.

The next step in the neural network is to "activate" the neurons by applying some nonlinear function that acts upon certain thresholds. This is analogous to how hedge funds determine the size of their trading positions. In the trading language, such a neuron-activation function is more commonly known as the "reaction" function. Following Bouchouev (2023), we use the reaction function $R(z)$ shown in Figure 2.

Figure 2
An Example of a Reaction Function



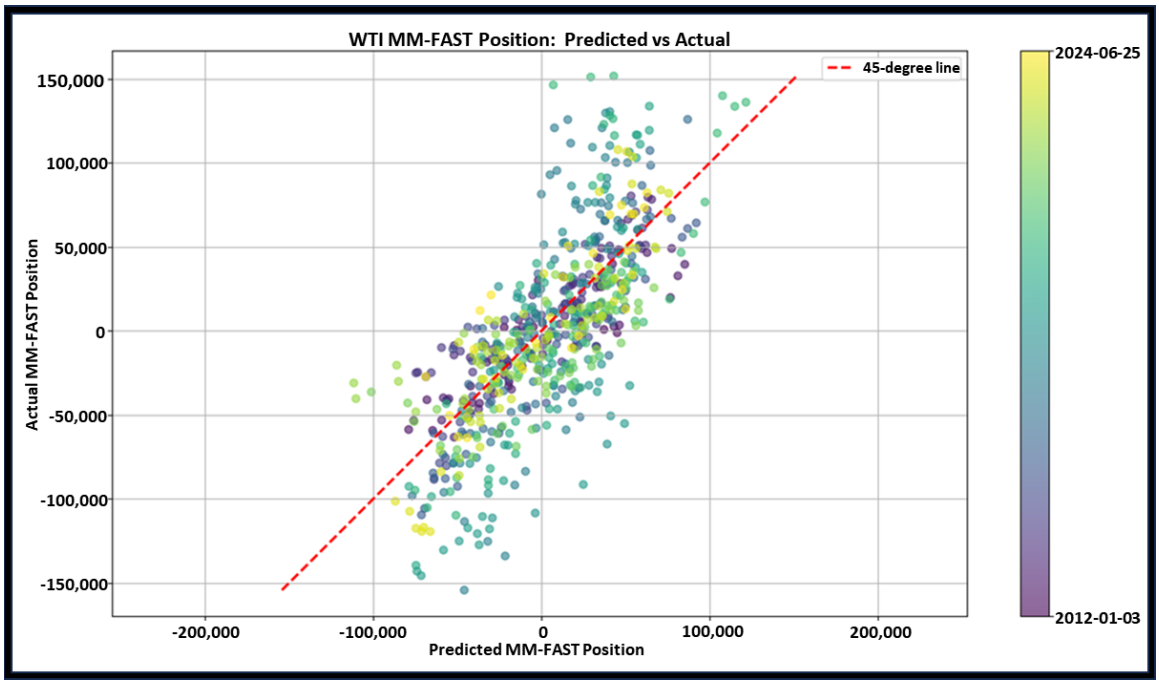
We apply the same reaction function to all three neurons. This function transforms the variable z, which represents a linear combination of momentum signals into the scale of $(-1, +1)$ and determines the size of the position as a percentage of the maximum budget allocation to the strategy.

The linear combination of transformed features produces the target forecast $\hat{y}_t$ for the normalized MM-FAST positions, where we use the hat notation to indicate that this variable is an estimate. The forecast $\hat{y}_t$ is then compared to the target y_t which represents the actual but also normalized futures positions held by MM-FAST. The network is trained by searching for incoming weights $w_{k,n}, k = 1,2,3, n = 10,20, \dots ,250$

and outgoing weights $W_k, k = 1,2,3$, which jointly minimize the difference between the forecast $\hat{y}_t$ and the target y_t. To improve the stability of the process, we apply a regularization technique to the minimizing cost function.

Figure 3 shows the results for the WTI market, which compare MM-FAST positions generated by our neural network to their actual positions.

Figure 3
Predicted Versus Actual MM-FAST Positions in the WTI Market



The model, based on this simple neural-network framework, produces a relatively high R^2 of 0.56, meaning that approximately half of the variance of hedge fund positions can be explained with technical, price-based momentum indicators. The results for Brent are remarkably similar, confirming the robustness of the method. This motivates our interpretation of the model-captured component as systematic or quantitative trading, with the residuals representing discretionary trading. In the main text of the paper we use this approach to separate positions held by systematic and discretionary traders and further analyze the time-series dynamics of their model-calculated positions.

Does the Behavior of Speculators Predict Futures Prices?

In keeping with the practical approach of this study, the metric we use for price forecasting is the economic performance of a trading strategy that buys and sells futures based solely on past positions held by speculators. Specifically, we analyze a popular investment strategy known as "following the flow" or "following the smart money."[4] This strategy assumes that hedge funds represent the "smart money," and that other investors may benefit by mimicking – as closely as possible – the actual positions held by these

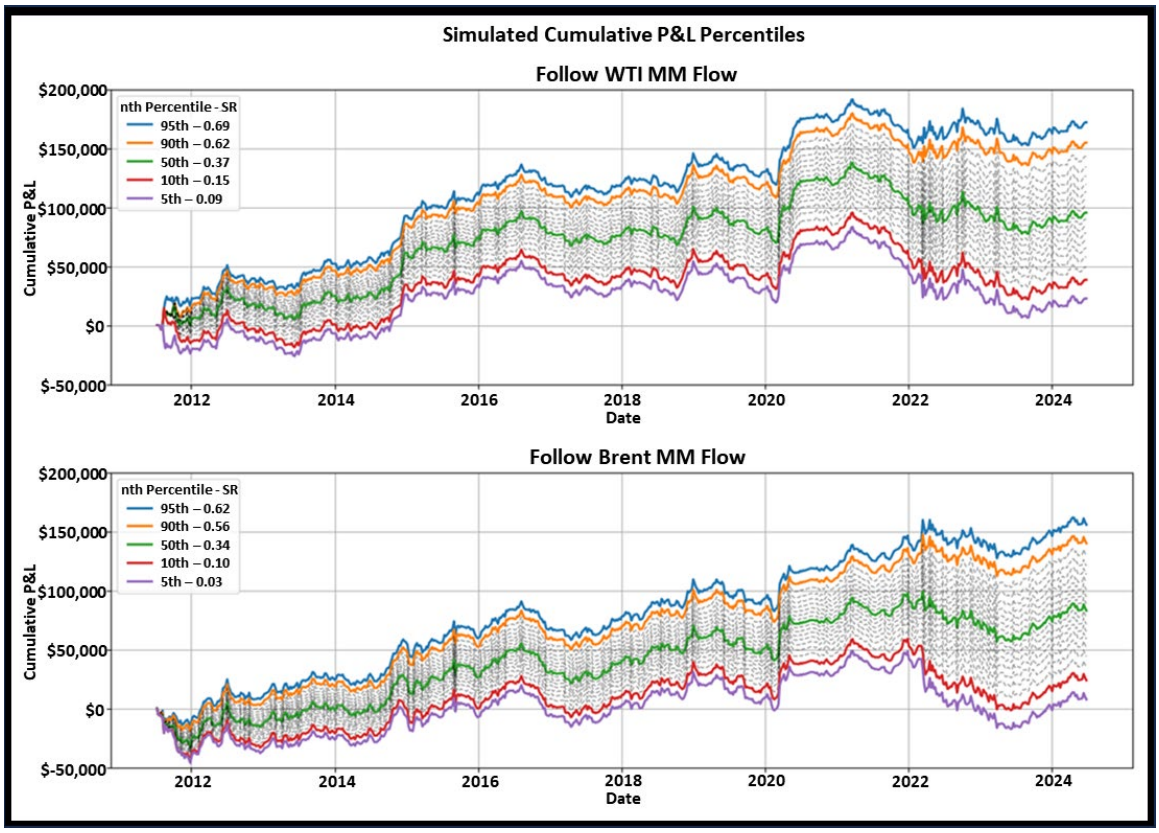
informed participants. In other words, investors would buy or sell oil futures as soon as it becomes known that the "smart money" had previously bought or sold contracts.

Having already observed the importance of moving averages of prices in explaining the behavior of "smart money," we now reverse the logic – following standard industry practices – and apply the crossover of moving averages to the actual positions held by MMs to assess whether they can generate price-forecasting signals.

However, the strategies based on moving averages are highly sensitive to the choice of the lookback periods, as detailed in the main text. Instead of trying to find the optimal parameters, we identify the region of acceptable parameters and randomize the set of parameters that define the trading signal in each period. We then run 1,000,000 simulations of this strategy using randomized parameters and group cumulative P&Ls by percentiles with 5% increments.

Figure 4 shows the cumulative performance of such strategies over time for WTI and Brent markets.

Figure 4
Cumulative P&L of "Follow-the-Flow" Strategy for MM Traders



It is evident that while these strategies performed well up until around 2021, their performance has since begun to deteriorate. This decline indicates that the information embedded in hedge fund positions has

become more widely recognized by other market participants, reducing the competitive advantage of simply mimicking their behavior. For additional results related to following and counter-following positions held by other market participants, we refer to the full text of the paper.

Conclusion

In this article, we have addressed several important topics related to speculation in the oil market.

First, we identified the main participants in oil futures trading and explained how their identities are often misunderstood or mislabeled in commonly used regulatory reports. This mischaracterization makes it particularly difficult to analyze oil speculation within the framework commonly applied to broader commodity portfolios. The nature of participants in the oil market is notably different – driven in part by the existence of a particularly large OTC market – which makes the oil market somewhat unique. This finding is consistent with previous research showing that the trading patterns of oil hedgers and speculators differ from those observed in many other commodities.

We then analyzed the actual profitability of key market participants and identified the so-called "fast money" group of traders as the primary winners in oil futures trading. This group is predominantly associated with quantitative hedge funds, many of which rely on technical models based on moving average crossovers. Using this concept, we constructed a simple neural-network algorithm. In our model, historical futures prices serve as features, and the hedge funds' position-sizing method – known as the reaction function – acts as the neural activation function in a machine learning framework.

The model explains approximately half of the variance in speculative positions held by hedge funds. The remaining variance is attributed to discretionary speculators and macro traders. We believe that this neural-network framework can be further improved by incorporating additional macroeconomic features, and we view this as a promising direction for future research.

In the final part of the article, we reverse the direction of causality between our two primary variables: oil prices and speculative positions. After explaining speculative positions using past prices, we then examine whether future prices can be predicted based on current and historical positions held by speculators. To study this question, we take a market-based approach and analyze the performance of a popular speculative strategy – simply following the hedge fund flows. While this strategy was profitable in the past, its performance has declined in recent years as the approach became more widely recognized by a broader set of market participants.

In short, speculative oil trading is always evolving – and the winners are those who evolve the fastest.

Endnotes

1 See Keynes (1923), Keynes (1930), and Hicks (1939).

2 In the conventional definition of the Sharpe ratio, one uses returns and subtracts the risk-free interest rate in the numerator. However, since futures positions do not require any initial cash investment besides a fairly small initial margin, one can ignore the contribution of the risk-free interest rate here.

3 The momentum signal $M_t(n)$ is defined as $M_t(n) = F_t - MA_t(n)$ where F_t is futures price and $MA_t(n)$ is n-day moving average. See the main text for more details.

4 See Bouchouev (2023), Chapter 6, for more details.

References

Bouchouev, I., 2023, Virtual Barrels: Quantitative Trading in the Oil Market, Cham, Switzerland: Springer.

Hicks, J. R., 1939, Value and Capital: An Inquiry into Some Fundamental Principles of Economic Theory, Oxford: Oxford University Press.

Kang, W., Rouwenhorst, K. G. and K. Tang, 2020, "A Tale of Two Premiums: The Role of Hedgers and Speculators in Commodity Futures Markets," *The Journal of Finance*, Vol. 75, No. 1, February, pp. 377-417.

Keynes, J. M., 1923, "Some Aspects of Commodity Markets," *The Manchester Guardian Commercial, Reconstruction Supplement*, March 29.

Keynes, J. M., 1930, A Treatise on Money (Vol. II), London: Macmillan.

Author Biographies

ILIA BOUCHOUEV, Ph.D.
Managing Partner, Pentathlon Investments LLC, U.S.A.

Dr. Ilia Bouchouev is the managing partner at Pentathlon Investments, an adjunct professor at New York University, and a senior research fellow at the Oxford Institute for Energy Studies. He is the former president of Koch Global Partners, where he launched and managed the company's global derivatives trading business for over 20 years. During his tenure, he introduced several energy derivatives products, and is widely recognized as one of the pioneers of energy options trading. Dr. Bouchouev holds a Ph.D. in Applied Mathematics, and his areas of expertise include commodity trading, energy economics, and option pricing. He is the author of "Virtual Barrels," a book on quantitative oil trading, which was published by Springer in 2023 and was named among the top ten quantitative books of the year: https://link.springer.com/book/10.1007/978-3-031-36151-7.

WU-YEN SUN
Visiting Researcher, King Abdullah Petroleum Studies and Research Center (KAPSARC), KSA

Wu-Yen (Jonathan) Sun is a visiting researcher in the Oil and Gas Program at KAPSARC, specializing in financial derivatives pricing and the financial sector's influence on crude oil prices. He holds an M.S. degree from New York University and has a diverse background in energy and finance. Before joining KAPSARC, he founded a solar power plant construction company and worked as a consultant at KPMG, focusing on solar and offshore wind power projects. His international experience in renewable energy development and market analysis supports KAPSARC's mission to advance global energy solutions.

AI Won't Take Our Job: It Will Reshape It

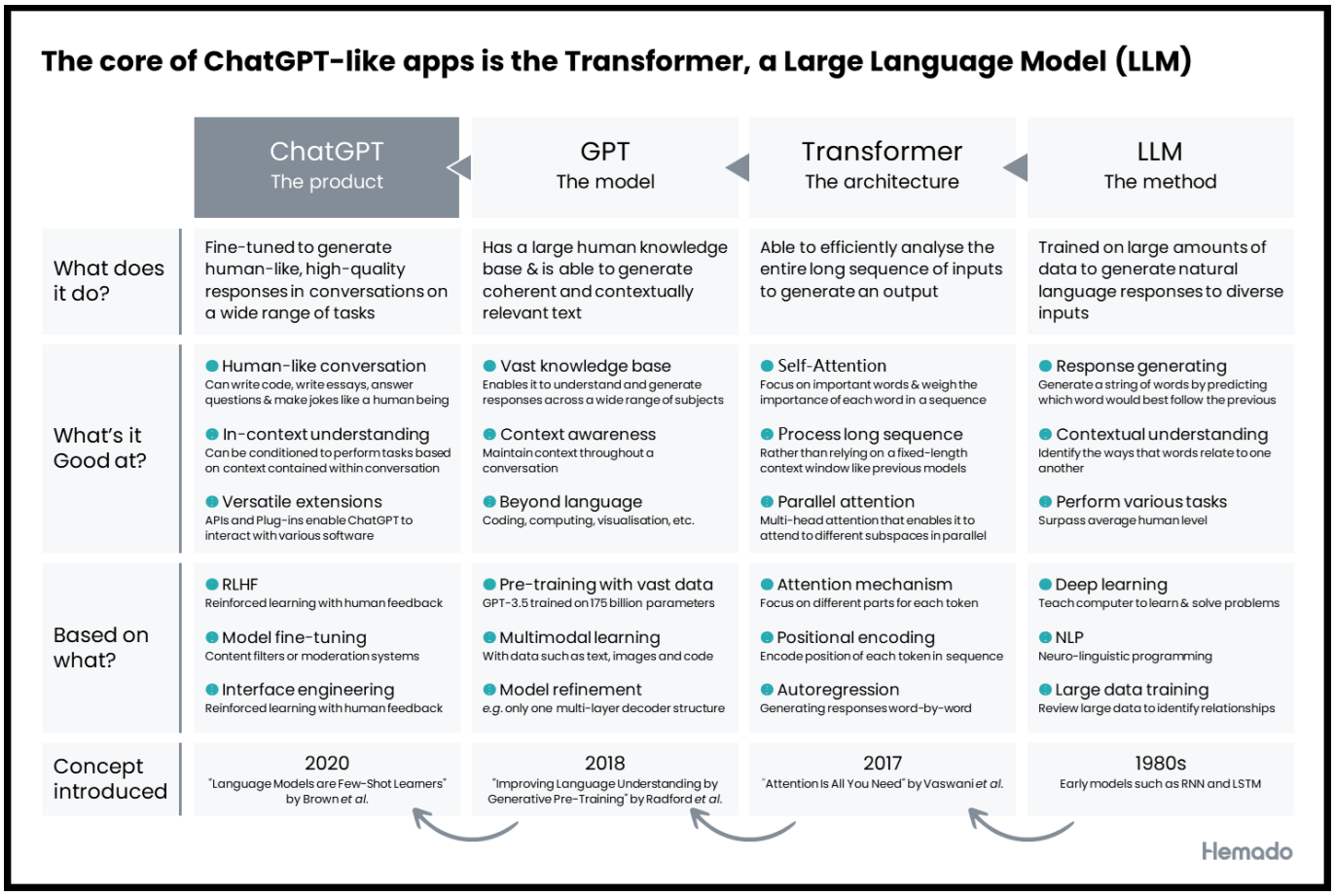
Victor Liu, Ph.D.
Managing Director, Hemado Green Energy, U.K.; and Editorial Board Member, *Commodity Insights Digest*

At Hemado, an energy consultancy, our exploration and integration of artificial intelligence (AI) has been an enlightening journey. In this digest article, I will share insights from this experience, discuss the impact of AI on research and consulting roles, and suggest some AI applications that we have found both useful and user-friendly.

We Do Not Need Many Chatbots

In March 2023, Bloomberg unveiled BloombergGPT, a powerful language model built specifically for the financial sector. But, unlike ChatGPT, BloombergGPT does not offer a public-facing chatbot interface, instead focusing on improving the efficiency of data retrieval for its users. Why didn't Bloomberg opt for an omniscient financial chatbot like ChatGPT? Additionally, why haven't there been significant advancements in the corporate adoption of GPT-like products since BloombergGPT's launch?

Figure 1



The answer lies in the rigorous performance standards required in high-stakes sectors such as energy and finance. Creating a specialized large language model (LLM) for a particular industry may seem like a good idea, but it is not that simple. It is hard to match the performance of generic models such as ChatGPT, Claude and Bard with a custom LLM. Because our institutional clients have seen powerful general-purpose models like ChatGPT, their expectations are extremely high. This can put specialized language models in a difficult position when we showcase them.

In addition, the implementation of AI in specialized areas still faces significant hurdles. While AI can speed up processes in low-hanging fruit areas such as data retrieval and generation, its effectiveness in specialized operations remains questionable. For example, the inconsistency inherent in generative AI stands in stark contrast to earlier AI products that offered greater precision. While specialized AIs such as facial recognition might have an accuracy rate of 99.999%, the precision of LLMs is less assured and the product could reveal its limitations soon after stunning at first.

Besides these issues, firms find it hard to track how their employees use generative AI. There's a risk of confidential data being uploaded and leaking in future talks. Earlier this year, Samsung limited its employees' use of ChatGPT following reports that engineers shared source code with the chatbot.

Figure 2

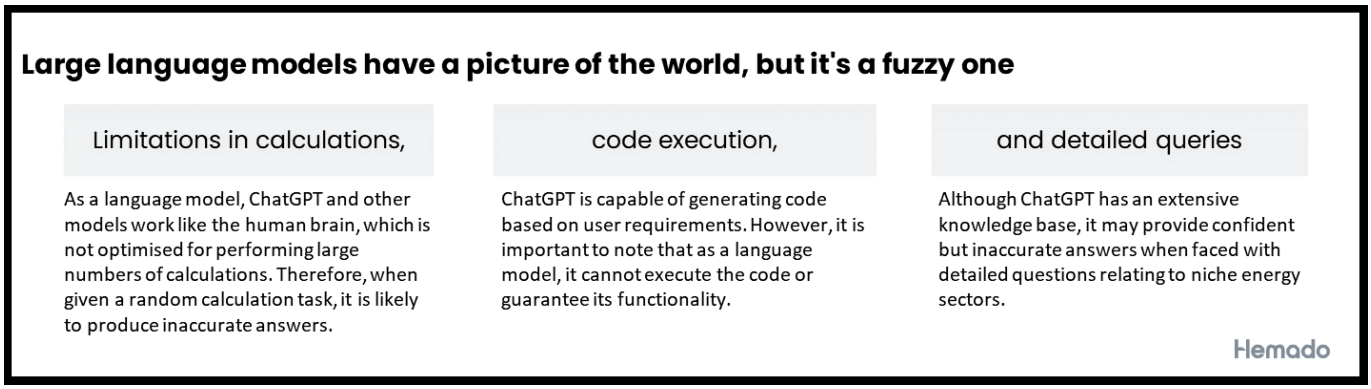


Figure 3

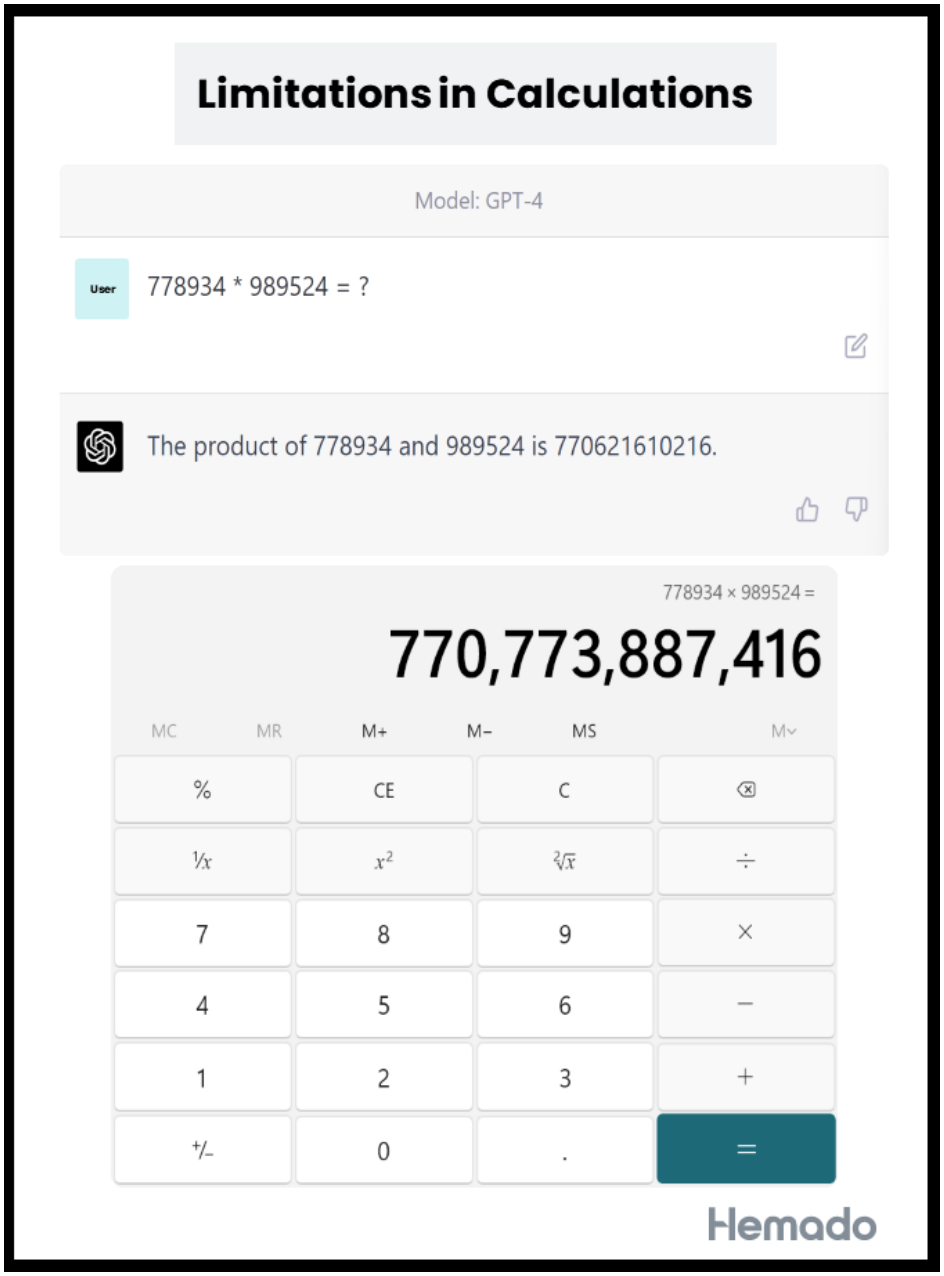

Limitations in Calculations
Model: GPT-4
User
778934 * 989524 = ?
The product of 778934 and 989524 is 770621610216.
778934 × 989524 =
770,773,887,416
MC MR M+ M− MS M⌄
% CE C
1/x x² ∛x ÷
7 8 9 ×
4 5 6 −
1 2 3 +
+/− 0 . =
Hemado

Figure 4

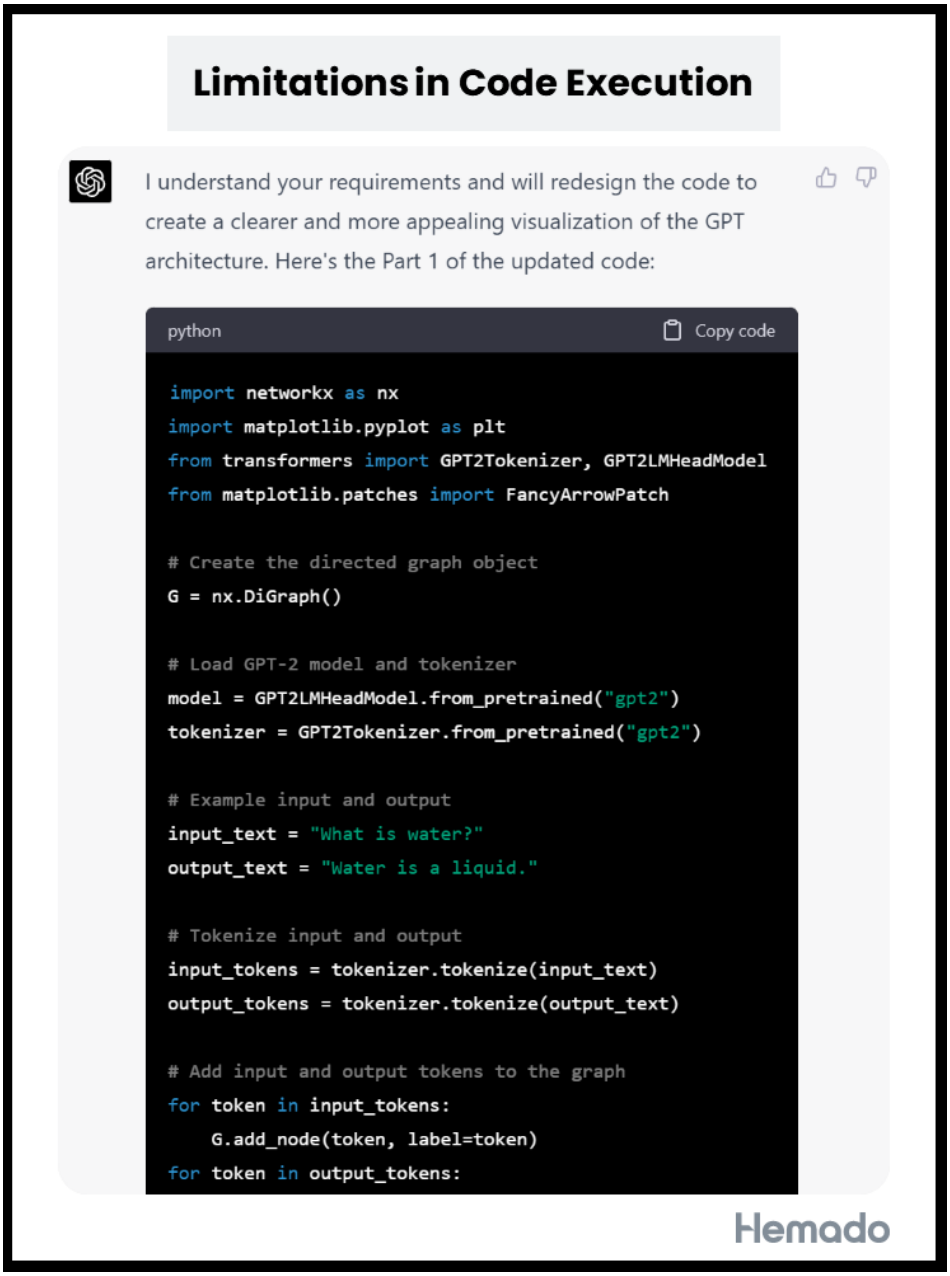
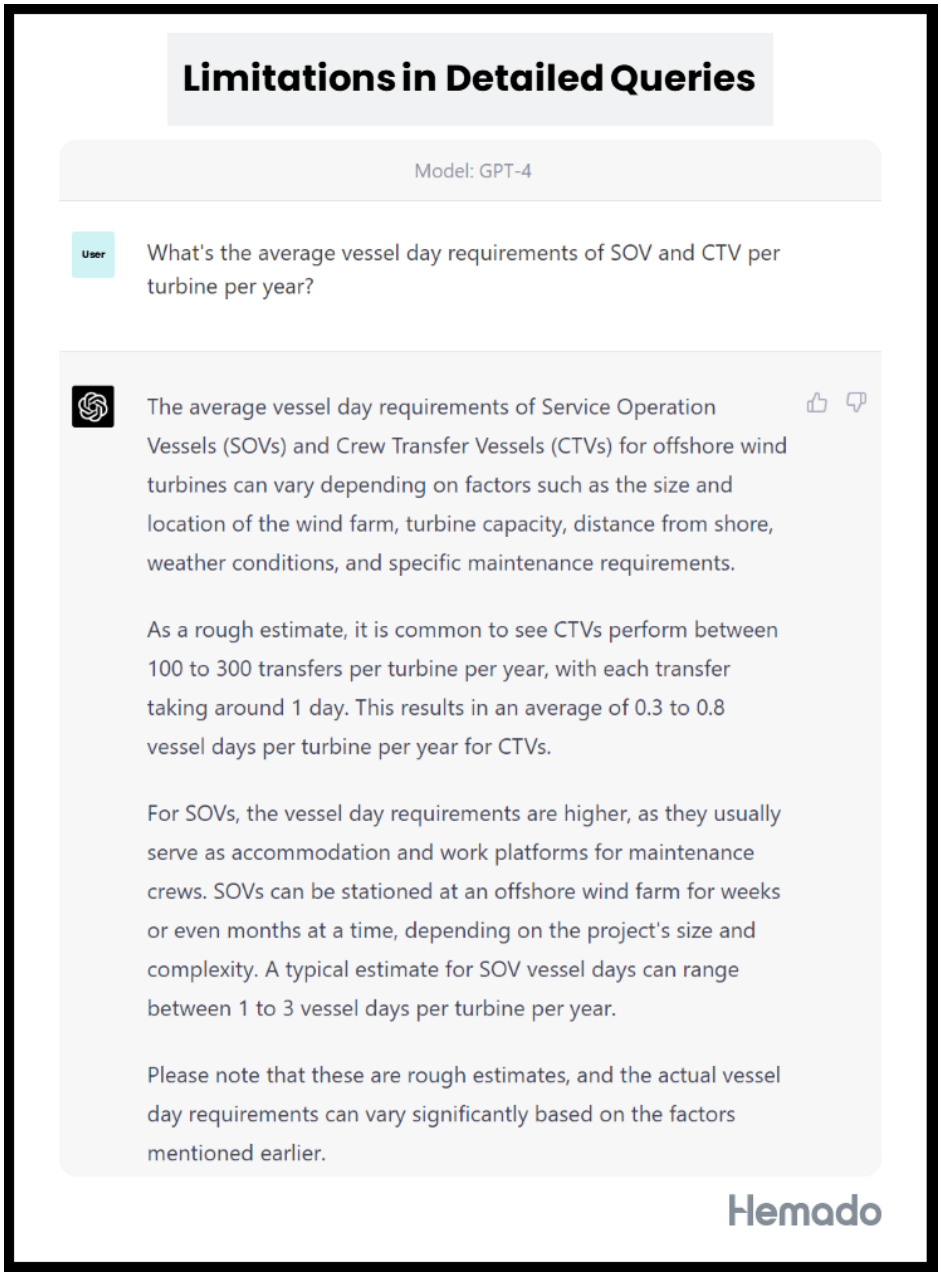

Limitations in Code Execution
I understand your requirements and will redesign the code to create a clearer and more appealing visualization of the GPT architecture. Here's the Part 1 of the updated code:
python
Copy code
import networkx as nx
import matplotlib.pyplot as plt
from transformers import GPT2Tokenizer, GPT2LMHeadModel
from matplotlib.patches import FancyArrowPatch
# Create the directed graph object
G = nx.DiGraph()
# Load GPT-2 model and tokenizer
model = GPT2LMHeadModel.from_pretrained("gpt2")
tokenizer = GPT2Tokenizer.from_pretrained("gpt2")
# Example input and output
input_text = "What is water?"
output_text = "Water is a liquid."
# Tokenize input and output
input_tokens = tokenizer.tokenize(input_text)
output_tokens = tokenizer.tokenize(output_text)
# Add input and output tokens to the graph
for token in input_tokens:
G.add_node(token, label=token)
for token in output_tokens:
Hemado

Figure 5



As a result, high-stakes industries that have zero tolerance for "hallucination," *i.e.*, the confident delivery of incorrect information, often find language models unsuitable. This often leads to the rejection of AI implementations at the proof of concept (POC) stage.

At Hemado, our attempt to develop a GPT-like product for renewable energy has resulted in more accurate, structured answers in certain cases. However, it is not immune to hallucination and its flexibility is not on a par with general models. As a result, the value proposition as a standalone chatbot was unclear to our customers, given their high demands for reliability and structured delivery. So, like Bloomberg, we

decided to use our model to improve our internal processes, leveraging the strengths of AI while mitigating its limitations.

AI is a Tool, Not a Replacement

The truth is that our customers do not just want a fancy AI model. They want a simple, reliable solution that solves their specific problems. Even if we can deliver more powerful, unified AI models in the near future, delivering a solution that really works for our clients is still a big challenge.

That is why the primary goal of using AI in research and consulting is not just to demonstrate AI capabilities, but to deliver robust, professional end-user solutions. Delivering AI directly to clients may not be the best approach. The benefits that specialized AI models, such as BloombergGPT as a chatbot, offer over more general models, such as ChatGPT or Bard, may be marginal and insufficient to simplify client tasks.

Instead, researchers and consultants who integrate AI to improve their productivity, quality and presentation are likely to achieve more fruitful results. However, they must take full responsibility for managing the potential risks associated with the use of AI.

Either way, if your job involves high stakes and precise decision-making, AI is unlikely to usurp your role: not any time soon.

AI Literacy is the Way Forward

Jensen Huang, CEO of Nvidia, likened the emergence of ChatGPT to the "iPhone moment" of AI. However, the full potential of this "moment" may still be ahead of us. Just as the iPhone did not fully mature until the release of iPhone 4, today's AI applications are still in their infancy, comparable to the first basic iPhone Operating System (IOS) apps such as Calculator and Flashlight.

The future of AI is likely to see advances in multimodal capabilities and longer context handling. This will provide a robust platform for AI applications to really shine. So it is important not to rush, but to prepare and position ourselves for the exciting era of AI applications yet to come.

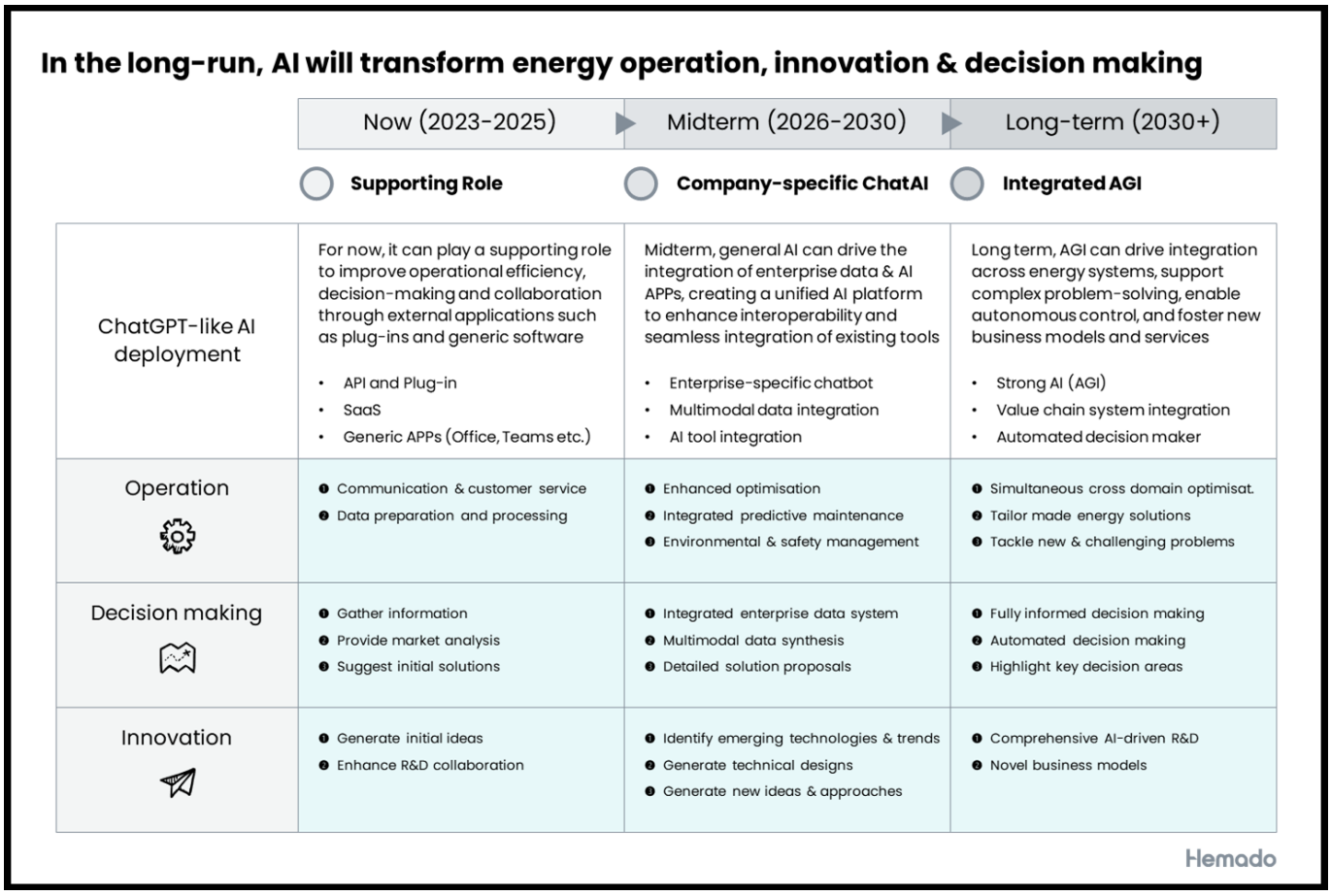
For the time being, I think it is better to think of AI as a tool to enhance our work, rather than an extension of our business. What we can do is integrate AI seamlessly into our workflow. AI can dramatically improve our productivity. It can handle routine tasks such as translation, proofreading and summarization. It can also be a tireless knowledge tutor, helping you quickly grasp the essence of new topics or fields. It can even help with initial idea generation and design, providing an initial integration of a framework on which to build.

Many AI tools have emerged that can significantly improve the way we conduct research, analyze data and present ideas. For example, some recommended tools include:

- Audiopen.ai for instant voice-to-text transcription.
- Rewind.ai as AI assistants to answer any questions based on your past interactions with your computer.
- Readwise.io/read for serious readers, providing summaries and Q&A for saved articles.
- GPTs as specialized AI agents for specific tasks.
- Perplexity AI for online searches with accurate and informative answers.
- DeepL translator for translating texts between different languages.

Anyway, the key is to use AI applications as a tool within your workflow, taking what AI can offer and enhancing it to increase your productivity and meet the professional standards of your field.

Figure 6



In conclusion, while the future promises an AI revolution, change is often incremental. In the words of Roy Amara, the American scientist and futurist, "We tend to overestimate the impact of a technology in the short term and underestimate the impact in the long term."

"AI won't replace your job, but a person using AI will." As AI evolves, it will redefine our roles. Efficiency gaps will widen, the nature of work will change and standards of professionalism will rise. Those who embrace AI at the right pace, with a clear purpose, will add increasing value along the way.

References

Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J., Winter, C., Hesse, C., Chen, M., Sigler, E., Litwin, M., Gray, S., Chess, B., Clark, J., Berner, C., McCandlish, S., Radford, A., Sutskever, I. and D. Amodei, 2020, "Language Models are Few-Shot Learners," *Advances in Neural Information Processing Systems,* Vol. 33, pp. 1877-1901.
Accessed via website: https://arxiv.org/pdf/2005.14165.pdf on November 21, 2023.

Radford, A. and K. Narasimhan, 2018, "Improving Language Understanding by Generative Pre-Training," OpenAI Technical Report, June 11.
Accessed via website: https://www.semanticscholar.org/paper/Improving-Language-Understanding-by-Generative-Radford-Narasimhan/cd18800a0fe0b668a1cc19f2ec95b5003d0a5035 on November 21, 2023.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L. and I. Polosukhin, 2017, "Attention is All You Need," NIPS'17: Proceedings of the 31st International Conference on Neural Information Processing Systems, December, pp. 6000-6010.
Accessed via website: https://dl.acm.org/doi/10.5555/3295222.3295349 on November 21, 2023.

Author Biography

VICTOR LIU, Ph.D.
Managing Director, Hemado Green Energy, U.K.

Dr. Victor Liu is an energy consultant with more than 15 years of global experience. He was previously Director of Consulting at Wood Mackenzie and Head of Transaction Services at Infield Systems in London.

Dr. Liu has extensive knowledge and experience in renewables and energy transition. At Hemado, he has advised international energy companies on renewable energy, digitalization strategy, market expansion plans and investment appraisal in cutting-edge green technologies such as hydrogen, carbon capture, utilization and storage (CCUS), floating wind, microgrid and virtual power plants (VPP).

He has also led several due diligence projects for mergers-and-acquisitions (M&A) transactions on behalf of top-tier financial institutions, including KKR, Goldman Sachs, Carlyle, Basalt and Alchemy. He has also led strategic or commercial advisory projects for a number of industrial companies including General Dynamics, CIMC, PetroChina, Nippon Steel, Acteon, Hydrasun and Centurion.

Dr. Liu is a regular speaker at global energy conferences and has published his findings in academic and industry journals. He is also a member of the editorial board of Bayes Business School's *Commodity Insights Digest* (*CID*). His academic qualifications include a Ph.D. in Financial Engineering from Bayes Business School (formerly Cass) in London, and a B.Sc. and M.Sc. (first class) in Economics from the University of Bath, U.K.

Mitigating AI Infrastructure Risks

Jodie Gunzberg, CFA
Managing Partner, InFi Strategies, U.S.A. and Editorial Board Member, *Commodity Insights Digest*

Hilary Till
Managing Co-Editor, *Commodity Insights Digest*; Honorary Senior Visiting Fellow, Bayes Business School, City St George's, University of London, U.K.; and Principal, Chicago-based Premia Research LLC, U.S.A.

Introduction

The rapid expansion of artificial intelligence (AI) infrastructure is being heralded as the next industrial revolution, with investments expected to surpass $1 trillion over the next decade (Dell'Oro Group, 2025). From data centers and semiconductor fabrication to power grid enhancements and software development, the scale of capital deployment is unprecedented. However, history has shown that large-scale infrastructure investments are not immune to obsolescence. Technological advancements, regulatory shifts, and market miscalculations can turn once-promising investments into stranded assets.

This article examines the risks associated with AI infrastructure investments, focusing on energy consumption, natural resource constraints, and historical parallels from past industrial buildouts. It explores factors that could lead to infrastructure obsolescence, market misalignment, and the ways that investment strategies may evolve in response to AI's future.

Explosive AI Infrastructure Growth May Pose Investment Risks

The rapid expansion of AI infrastructure represents an unprecedented investment surge. However, in Appendix A, *AI "Stranded Assets?"*, we caution that this aggressive buildout may be a risky bet on existing technological frameworks. Rapid advancements in computing efficiency and emerging alternative models could render significant portions of this investment obsolete. Regulatory shifts and mounting energy constraints could further threaten the long-term viability of AI infrastructure.

Much of this investment is concentrated in hardware, with companies such as NVIDIA, Intel, and AMD driving semiconductor advancements tailored for AI workloads. The global expansion of AI-focused data centers is accelerating, led by cloud providers such as Amazon Web Services, Microsoft Azure, and Google Cloud. The International Energy Agency (IEA, 2024) projects that AI-related electricity demand could increase by 50% by 2030, requiring significant power grid upgrades. Additionally, the Semiconductor Industry Association estimates that over $500 billion will be allocated to new semiconductor manufacturing plants to meet AI's growing processing needs (Spiceworks, 2025). While these investments aim to support AI's transformative potential, they carry the risk of premature obsolescence if technological advancements outpace current infrastructure.

Obstacles to AI's Infrastructure Development

AI computing workloads, particularly large-scale model training and inference, require immense amounts of energy. Training large-scale AI models like GPT-4 requires substantial energy, with estimates suggesting that its electricity consumption could be up to 50 times higher than previous-generation models like GPT-3, which itself used approximately as much electricity as 130 U.S. households in a year (WEF, 2024). If demand continues to escalate, AI-related power consumption could place a significant burden on global electricity grids. By 2030, AI is projected to account for nearly 10% of global electricity usage, up from 4% estimated today, which could strain energy networks and increase operational costs for AI firms (EPRI, 2024). The concentration of AI-driven power demand in technology hubs such as Silicon Valley, Singapore, and Ireland exacerbates this issue, as local grids in these regions struggle to keep pace.

Jodie Gunzberg, CFA, Chairperson of the CFA NJ Group and former board member of the CFA Society in New York and co-author of this article, celebrated CFA Society of NY's Veterans Roundtable by ringing the NASDAQ Stock Market Closing Bell in New York City. Gunzberg is a Managing Partner of InFi Strategies and an Editorial Board member of the *CID*.

To mitigate AI's growing energy footprint, companies like Google and Microsoft have committed to transitioning their AI operations to renewable energy. However, reliance on renewables presents challenges, as wind and solar power depend on weather conditions, leading to fluctuations in availability. AI data centers require consistent uptime, and without adequate energy storage solutions, they may still need to rely on fossil fuel backups. The modernization of power grids is another significant hurdle, as existing infrastructure may not be equipped to handle AI's high-intensity, fluctuating energy demands. The cost of upgrading these grids is substantial, and investments may not align with AI's rapid innovation cycles. Additionally, competition for renewable energy resources among AI firms, manufacturers, and electric vehicle companies could create supply constraints, driving up costs and limiting access for new entrants.

Policymakers and energy providers will need to develop strategies to balance demand and ensure equitable access to sustainable power sources.

Beyond energy concerns, AI infrastructure depends on a fragile global supply chain. Advanced AI chips require minerals that are primarily sourced from a few geopolitical hotspots. Over 50% of the world's cobalt reserves are from the Democratic Republic of Congo (DRC), where mining conditions raise serious ethical concerns, including child labor and environmental degradation (USDOL, 2024; USGS, 2024). Similarly, China dominates the supply chain of the neodymium magnets used in AI infrastructure, including cooling systems and power management units, as highlighted by Paterson (2023) and USDOC (2022). Supply chain disruptions, whether caused by political conflict, export restrictions, or resource depletion, could significantly hinder AI hardware production and investment.

The extraction of these materials carries significant environmental costs. Lithium mining, crucial for AI hardware batteries and energy storage systems, consumes vast amounts of water, depleting already scarce resources in regions such as Argentina, Bolivia, and Chile. In South America's "Lithium Triangle," lithium extraction requires nearly 500,000 gallons of water per ton, leading to land degradation and conflicts with local communities over water rights (IER, 2020). Meanwhile, cobalt sourcing's overreliance on the DRC exposes AI chipmakers to supply shocks and regulatory scrutiny. Diversifying sources and investing in synthetic materials could mitigate these risks.

Geopolitical tensions further complicate AI's supply chain vulnerabilities. The ongoing U.S.-China trade war has led to semiconductor export controls and restrictions on critical materials like gallium and germanium (Lv and Munroe, 2024). In response, the U.S. CHIPS Act and the European Union's semiconductor sovereignty initiatives aim to reduce reliance on Asian manufacturers. However, these efforts require years to materialize, leaving the industry exposed to short-term supply disruptions. Another bottleneck is the high cost and long lead time associated with building semiconductor fabrication plants, which costs about $10 billion and takes three to five years and 6,000 construction workers to complete (Intel, 2023). Companies such as TSMC, Intel, and Samsung are expanding chip production, but shifting technological demands could render existing facilities obsolete before they become profitable. Recent global chip shortages, exacerbated by the COVID-19 pandemic, have further highlighted the fragility of AI hardware supply chains. Without investment in alternative semiconductor technologies like 3D chip stacking and nanomaterials, AI infrastructure could face severe limitations.

Ensuring long-term stability in AI infrastructure will require multipronged solutions, including diversified sourcing of rare earth minerals, stronger regulatory oversight for ethical mining practices, and advancements in semiconductor production. Public and private sector collaborations will be critical in building resilient supply chains that can withstand geopolitical and environmental challenges. Sustainable innovation, energy-efficient AI models, and breakthroughs in alternative energy sources will likely be essential to ensuring the long-term viability of infrastructure investments.

Historical Precedents of Stranded Infrastructure

History is filled with examples of large-scale infrastructure investments that became obsolete due to technological advancements, regulatory changes, or shifting market dynamics. These cases offer valuable lessons for AI infrastructure investments, which could face similar risks in the future.

In the 1970s, nuclear power was widely believed to be the future of global energy, prompting significant government and private sector investments in nuclear plants with the expectation of delivering cheap, abundant electricity. However, safety concerns, cost overruns, and increasing regulatory restrictions led to the premature abandonment of many projects. High-profile incidents such as the Three Mile Island disaster in 1979 and the Chernobyl explosion in 1986 heightened public opposition, causing many countries to scale back or phase out nuclear energy altogether, justifiably or not (NRC, 2022; Ahearne, 1987). At the same time, advancements in natural gas extraction, particularly through hydraulic fracturing (fracking), provided a cheaper and more flexible energy alternative, further diminishing the economic viability of nuclear plants until very recently.

The late 1990s saw a similar miscalculation with the overbuilding of fiber-optic networks during the dot-com boom. Telecom companies, anticipating explosive internet growth, laid thousands of miles of fiber-optic cables, expecting surging demand for broadband connectivity. However, demand failed to materialize at the expected rate, leaving much of the installed infrastructure unused and resulting in a glut of "dark fiber" that took years to become profitable (Hecht, 2016). A comparable risk exists for AI infrastructure today, as data centers and semiconductor fabrication plants are being constructed at an unprecedented pace without clear long-term demand projections.

A parallel can also be drawn to the U.S. railroad boom of the late 19th century, which was driven by government incentives and speculation. Many companies built redundant or poorly planned rail networks, leading to financial collapses when demand failed to meet expectations. The Panic of 1873 and the wave of railroad bankruptcies that followed underscored the dangers of infrastructure overinvestment without proper demand forecasting, as speculative expansion in railroads outpaced economic viability (Fogel, 1964; PBS, 2025). AI infrastructure investments, particularly in data centers and specialized AI chips, could face a similar fate if technological advancements render current approaches obsolete.

More recently, the overinvestment in coal-fired power plants in the early 2000s highlights how rapid technological and regulatory shifts can quickly change the viability of major energy investments. Governments and utilities worldwide invested billions in new coal plants, believing them essential for long-term energy security. However, advancements in renewable energy and tightening regulations on carbon emissions made many coal plants financially unviable. In the U.S., nearly 50% of the coal plants operating in 2010 have since been shut down or converted to natural gas, often well before their intended economic lifespan due to competition from cheaper and cleaner energy sources (EIA, 2019). This serves as a reminder of how quickly market conditions and regulatory policies can shift, rendering once-essential infrastructure obsolete.

These historical examples illustrate the risks of overinvestment in static infrastructure that fails to adapt to evolving technological or regulatory landscapes. With AI infrastructure requiring heavy investment in energy-intensive data centers and specialized semiconductors, flexibility and adaptability must remain central considerations. Policymakers, investors, and industry leaders must ensure that AI infrastructure investments align with shifting market trends and technological innovations to mitigate the risk of stranded assets.

Managing the Risks of AI Infrastructure Investments

AI infrastructure investments designed for adaptability may mitigate risks. One approach is to focus on scalability by developing modular data centers and edge computing architectures that reduce the risk of overinvestment in centralized hubs. Flexible infrastructure allows for incremental scaling rather than large, inflexible capital expenditures. Emerging trends in decentralized computing and federated learning could decrease reliance on large centralized facilities, minimizing the risk of stranded assets. Additionally, hybrid cloud strategies that balance on-premises, edge, and cloud computing resources can optimize cost and efficiency (Sakata, *et al*. 2024; Eriksvik, 2024).

Prioritizing energy efficiency is another key strategy in mitigating risks. AI firms can reduce power consumption by leveraging energy-efficient computing solutions such as photonic chips and neuromorphic processors. Optimizing AI model architectures through techniques like sparsity and low-power processing units can significantly lower energy demand. Governments and industry leaders are also exploring advanced cooling technologies, such as liquid and immersion cooling, to further optimize AI infrastructure. Policy incentives and carbon offset mechanisms could encourage firms to adopt greener computing solutions. Research suggests that improving energy efficiency across the AI model development cycle can substantially reduce carbon impact (Wu *et al*., 2024).

A diversified investment strategy can also help hedge against infrastructure obsolescence. While current AI models rely heavily on classical computing architectures, advancements in quantum computing, optical processing, and next-generation semiconductor technologies could disrupt the industry. Investing in a broader portfolio that includes these emerging technologies can mitigate risks associated with overreliance on current AI infrastructure models. Additionally, investments in complementary sectors such as sustainable energy and advanced materials provide further protection against infrastructure stagnation. Major firms like BlackRock and Microsoft have launched multi-billion-dollar funds focused on AI infrastructure, including data centers and energy projects, signaling a shift toward diversified AI investment strategies (Masters, 2024).

Conclusion

AI infrastructure investment represents a major economic shift, akin to past industrial revolutions. However, history shows the dangers of overinvestment in static technologies. By ensuring AI infrastructure is adaptable, sustainable, and responsive to technological advancements, investors and policymakers can mitigate the risk of stranded assets.

Mitigating AI Infrastructure Risks

The lessons from past industrial and technological transformations highlight the need for cautious optimism. While AI offers immense opportunities, the pace of innovation means that investments made today could become obsolete faster than anticipated. The financial consequences of misaligned investment strategies could be severe, impacting not only corporations but also broader market stability.

Furthermore, the global nature of AI development necessitates a coordinated approach between governments, industry leaders, and investors. Policymakers must strike a balance between fostering AI innovation and implementing regulations that prevent unsustainable energy consumption and environmental degradation. Collaboration between public and private entities will be essential to creating a resilient AI infrastructure that benefits society without leading to unnecessary economic losses.

Ultimately, those investing in AI infrastructure must remain flexible, embracing a forward-looking approach that incorporates the potential for new breakthroughs and shifts in market dynamics. By learning from past infrastructure investment missteps, stakeholders can navigate this new era of AI-driven transformation while minimizing the risks of stranded assets.

Appendix A: AI "Stranded Assets"?

Another Industrial Revolution

The consequences of the latest generation of Artificial Intelligence (AI) algorithms are undoubtedly staggering. We have arguably entered "a new era, just like Gutenberg's press and the invention of the steam engine revolutionized their respective eras," quoting Manthos (2022) on powerful technologies. The software architecture underpinning large language models (LLMs) is astonishingly efficient in effectively storing *all* manner of concepts that are accessible from the corpus of the internet. More technically, LLMs use "optimization algorithms … that iteratively adjust model parameters … in a way that brings the model's predictions closer to the original text [so that] the model 'learns' the linguistic patterns in its training data (and, by extension, the 'knowledge' conveyed in those linguistic patterns)," as explained by Bergmann (2024). These explanations summarize the significance of the current generation of AI models by emphasizing *what they effectively do*. Essentially, we are on the cusp of another Industrial Revolution where this time, instead of a skilled craftsman's work being mechanized, a large fraction of a knowledge worker's tasks will be automated.

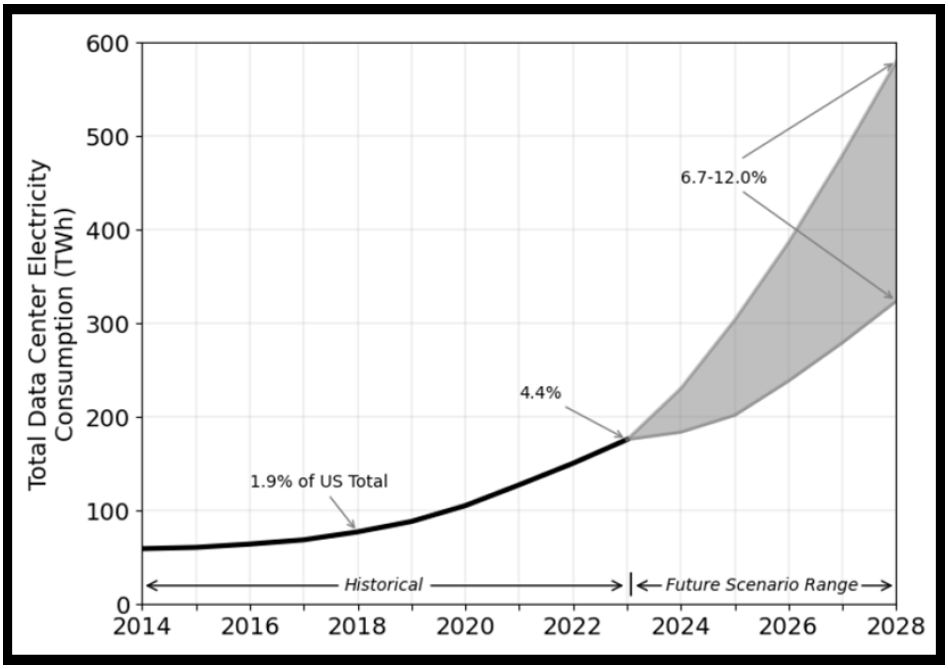
Hilary Till (LHS of photo), co-author of this article, moderated a panel discussion on "AI/ML in Finance – Shaping the Future of Financial Markets" at the London School of Economics (LSE) in an event co-sponsored by the LSE Data Science Institute and the International Association for Quantitative Finance (IAQF). The event took place on June 24, 2025. To Till's right are (L-to-R): Johannes Muhle-Karbe, Ph.D. (Imperial College London), Peter Hafez (RavenPack/BigData.com), Giovanni Beliossi (Axyon AI), and Petter Kolm, Ph.D. (New York University).

Huge and Costly AI Infrastructure Buildout

The consensus belief on how a company could stay on the cutting edge of AI was summarized in a memo by Karpathy (2018). He noted: "… progress in AI is primarily driven by systems – compute, data, infrastructure. The core algorithms we use today have remained largely unchanged ..." While allowing for

the possibility of improvements, the AI researcher later observed in Karpathy (2024) that since 2017, the current models have "proven remarkably resilient." Hence, the focus (for now) has been on a huge AI infrastructure buildout, which would cost *over $1-trillion* in the next several years alone and which would include spending on chips, data centers, the power grid, and other AI infrastructure, according to Covello (2024) and Nathan (2024). Relatedly, see Figure A-1 on the expected increase in U.S. data center electricity usage through 2028, "owing to the rapid emergence of AI hardware," as discussed in Shehabi *et al.* (2024).

Figure A-1
Total U.S. Data Center Electricity Use from 2014 through 2028



Source: Shehabi *et al.* (2024), Figure ES-1.

A Bet *Against* Innovation

One trillion dollars is a surprisingly large bet on there *not* being sufficient innovation in underlying AI principles, algorithms, and computer chips, and instead relying on enormous infrastructure expenditures. For example, Li (2023) compared current AI models to the field of physics before "Newton … started to write very simple laws in physics." As well, Liebreich (2024) predicted that the computer-chip "sector is moving so fast that there are almost certainly innovations … that could make AI compute dramatically more energy efficient."

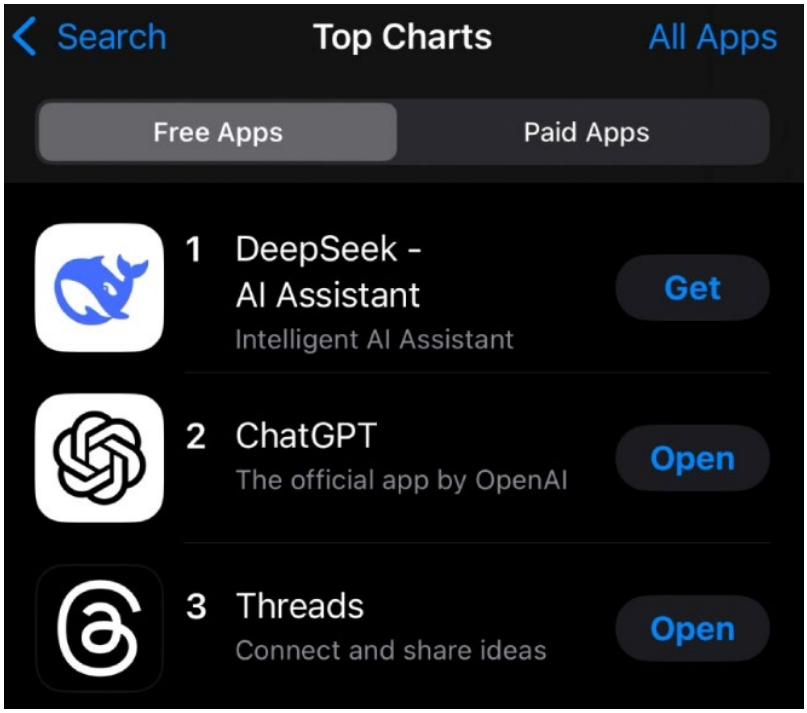
At some point, if there were sufficient breakthroughs, it is questionable whether the huge level of AI infrastructure investment would be needed. Angelo Sabatelle, who manages Moody's project finance team for the Americas, told Shi (2024): "You can't close your eyes to the fact that technology changes

quickly." Further, Covello (2024) warned: "[o]ver-building things the world doesn't have use for … typically ends badly."

Current Developments

Confirming the speed of innovation, the Chinese lab, DeepSeek-AI, has already created an open-source large language model whose "total training costs" were less than $6 million ("excluding the costs associated with prior research" and experimentation), according to DeepSeek-AI (2024). The Chinese lab's model outperforms other models "in accuracy ranging from complex problem-solving to math and coding," wrote Wu and Bosa (2025). In addition, DeepSeek relies on "inference-time computing [which means] they activate only the most relevant portions of their model for each query, and that saves money and computation power," explained Giuseppe Sette, President at AI market research firm Reflexivity, in Picchi (2025). These new developments, concluded Wu and Bosa (2025), have "called into question big tech's massive spend on building AI models and data centers," particularly as the market quickly endorsed DeepSeek-AI's efforts, as shown in Figure A-2. Responding to the launch of DeepSeek's models, the International Energy Agency questioned the power demands of AI, according to Moore *et al.* (2025).

Figure A-2
Top Charts in the Apple App Store for the iPhone on January 28, 2025



Source: Apple App Store.

Stranded Assets

These events remind us of a constant lesson in economic history, dating back to the nineteenth century. For example, we can look to the 1830s and 1840s when investors speculated on the building of a crucial canal that ultimately linked productive Illinois farmland to major population centers. The canal was successfully completed in 1848, by which time it was already on its way to becoming obsolete because of the opening of railways (Till, 2014). In other words, the canal essentially became a *stranded asset* by the time it was built.

Will this historical lesson be repeated?

References

Ahearne, J. F., 1987, "Nuclear Power after Chernobyl," *Science*, Vol. 236, No. 4802, May 8, pp. 673-679. Accessed via website: https://doi.org/10.1126/science.3576192 on February 13, 2025.

Bergmann, D., 2024, "What is Instruction Tuning," IBM, April 5. Accessed via website: https://www.ibm.com/think/topics/instruction-tuning#:~:text=Instruction%20tuning%20is%20a%20technique%20for%20fine-tuning%20large,labeled%20dataset%20of%20instructional%20prompts%20and%20corresponding%20outputs on March 23, 2025.

Covello, J., 2024, "Interview with Jim Covello," *Top of Mind*, Goldman Sachs Global Macro Research, Issue 129, June 25, pp. 10-11.

DeepSeek-AI, 2024, "DeepSeek-V3 Technical Report," December 27. Accessed via website: https://arxiv.org/pdf/2412.19437 on January 24, 2025.

Dell'Oro Group, 2025, "Data Center Capex to Surpass $1 Trillion by 2029," Press Release, February 5. Accessed via website: https://www.delloro.com/news/data-center-capex-to-surpass-1-trillion-by-2029/ on February 13, 2025.

EIA, 2019, "More Than Half of the U.S. Coal-Fired Electricity Generating Capacity Has Been Retired Since 2010," U.S. Energy Information Administration, July 26. Accessed via website: https://www.eia.gov/todayinenergy/detail.php?id=40212 on February 13, 2025.

EPRI, 2024, "Powering Intelligence: Analyzing Artificial Intelligence and Data Center Energy Consumption," Electric Power Research Institute, May 28. Accessed via website: https://www.epri.com/research/products/3002028905 on February 13, 2025.

Eriksvik, J., 2024, "Federated Machine Learning and Hybrid Infrastructure as Levers to Accelerate Artificial Intelligence," Algorithma. Accessed via website: https://www.algorithma.se/our-latest-thinking/federated-machine-learning-and-hybrid-infrastructure-as-levers-to-accelerate-artificial-intelligence on February 13, 2025.

Fogel, R., 1964, Railroads and American Economic Growth: Essays in Econometric History, Baltimore: Johns Hopkins University Press.

Hecht, J., 2016, "Boom, Bubble, Bust: The Fiber Optic Mania," *Optics and Photonics News*, October. Accessed via website: https://internethistory.org/wp-content/uploads/2020/01/OSA_Boom.Bubble.Bust_Fiber.Optic_.Mania_.pdf on February 13, 2025.

IEA, 2024, "Electricity 2024," International Energy Agency, January.
Accessed via website: https://www.iea.org/reports/electricity-2024 on February 13, 2025.

IER, 2020, "The Environmental Impact of Lithium Batteries," Institute for Energy Research, November 12.
Accessed via website: https://www.instituteforenergyresearch.org/renewable/the-environmental-impact-of-lithium-batteries/ on February 13, 2025.

Intel, 2023, "How a Semiconductor Factory Works," June 9.
Accessed via website: https://newsroom.intel.com/tech101/how-a-semiconductor-factory-works on February 13, 2025.

Karpathy, A., 2018, "Andrej Karpathy to Elon Musk," OpenAI email archives, January 31.
Accessed via website: https://openai.com/index/openai-elon-musk/ on January 24, 2025.

Karpathy, A., 2024, "Making AI Accessible with Andrej Karpathy and Stephanie Zhan," Sequoia Capital's AI Ascent Conference, March 20. Accessed via website: https://www.youtube.com/watch?v=c3b-JASoPi0 on January 24, 2025.

Li, F.-F., 2023, "Andrew Ng and Fei-Fei Li Discuss Human-Centered Artificial Intelligence," Stanford Online, May 11. Accessed via website: https://www.youtube.com/watch?v=UNhC6Ox0T0o on January 24, 2025.

Liebreich, M., 2024, "Generative AI – The Power and the Glory," *BloombergNEF*, December 24.
Accessed via website: https://about.bnef.com/blog/liebreich-generative-ai-the-power-and-the-glory/ on January 24, 2025.

Lv, A. and Munroe, T., 2024, "China Bans Export of Critical Minerals to U.S. as Trade Tensions Escalate," *Reuters*, December 3. Accessed via website: https://www.reuters.com/markets/commodities/china-bans-exports-gallium-germanium-antimony-us-2024-12-03/ on February 13, 2025.

Manthos, P., 2022, "From Aristotle's Logic to Boolean Algebra: How George Boole and Claude Shannon Created Modern Computers," *CodeX*, May 6. Accessed via website: https://medium.com/codex/from-aristotles-logic-to-boolean-algebra-how-george-boole-and-claude-shannon-created-modern-aeedb34e276b on January 24, 2025.

Masters, B., 2024, "BlackRock and Microsoft Plan $30bn Fund to Invest in AI Infrastructure," *Financial Times*, September 17. Accessed via website: https://www.ft.com/content/4441114b-a105-439c-949b-1e7f81517deb on February 13, 2025.

Moore, M., Johnston, I. and L. Pitel, 2025, "DeepSeek Threat Exposes Guesswork on AI Power Demand, Says IEA," *Financial Times*, January 28. Accessed via website: https://www.ft.com/content/0cc897c2-e12d-4143-81ff-e56c5381a5a1 on January 28, 2025.

Nathan, A., 2024, "Gen AI: Too Much Spend, Too Little Benefit," *Top of Mind*, Goldman Sachs Global Macro Research, Issue 129, June 25, p. 3.

NRC, 2022, "Backgrounder on the Three Mile Island Accident," U.S. Nuclear Regulatory Commission, April. Accessed via website: https://www.nrc.gov/reading-rm/doc-collections/fact-sheets/3mile-isle.html on February 13, 2025.

Paterson, S., 2023, "U.S. Tries to Wean Off China's Grip on Rare-Earth Magnets," Hinrich Foundation, January 25. Accessed via website: https://www.hinrichfoundation.com/research/article/tech/dependent-on-china-rare-earth-magnets/ on February 18, 2025.

PBS, 2025, "The Panic of 1873," PBS American Experience. Accessed via website: https://www.pbs.org/wgbh/americanexperience/features/grant-panic/ on February 13, 2025.

Picchi, A., 2025, "What is DeepSeek, and Why is it Causing Nvidia and Other Stocks to Slump?", CBS News MoneyWatch, January 25. Accessed via website: https://www.cbsnews.com/news/what-is-deepseek-ai-china-stock-nvidia-nvda-asml/ on January 28, 2025.

Sakata, E., Chen, C. and I. Yang, 2024, "Decentralized Collaborative AI with Federated Learning in Trustworthy Environment [S62427]," NVIDIA, March 18. Accessed via website: https://www.nvidia.com/en-us/on-demand/session/gtc24-s62149/ on February 13, 2025.

Shehabi, A., Smith, S., Hubbard, A., Newkirk, A., Lei, N., Siddik, M., Holecek, B., Koomey, J., Masanet, E. and D. Sartor, 2024, "2024 United States Data Center Energy Usage Report," Energy Analysis and Environmental Impacts Division, Lawrence Berkeley National Laboratory, December. Accessed via website: https://www.energy.gov/articles/doe-releases-new-report-evaluating-increase-electricity-demand-data-centers on January 24, 2025.

Shi, M., 2024, "Sticking Points in the AI Boom," *PitchBook News*, December 20. Accessed via website: https://pitchbook.com/news/articles/sticking-points-in-the-ai-boom on January 24, 2025.

Spiceworks, 2025, "Global Investment on 84 New Chip Fabs to Touch $500B by 2024," January 7. Accessed via website: https://www.spiceworks.com/tech/hardware/news/global-semiconductor-manufacturing-boom/#:~:text=The%20semiconductor%20industry%20association%20SEMI%20estimated%20chipmakers,28%20new%20plants%20will%20start%20in%202023 on February 13, 2025.

Till, H., 2014, "Futures Industry: A Story of Crisis and Opportunity," *Futures Magazine*, Cover Story, February, pp. 14-21. Accessed via website: https://www.premiacap.com/publications/Futures%20Magazine%20cover%20story%20Feb%202014.pdf on January 24, 2025.

USDOC, 2022, "The Effect of Imports of Neodymium-Iron-Boron (NdFeB) Permanent Magnets on the National Security," U.S. Department of Commerce, Bureau of Industry and Security, Office of Technology Evaluation. Accessed via website: https://www.bis.doc.gov/index.php/documents/section-232-investigations/3141-report-1/file on February 18, 2025.

USDOL, 2024, "Combatting Child Labor in the Democratic Republic of the Congo's Cobalt Industry (COTECCO)," U.S. Department of Labor. Accessed via website: https://www.dol.gov/agencies/ilab/combatting-child-labor-democratic-republic-congos-cobalt-industry-cotecco on February 13, 2025.

USGS, 2024, "Cobalt," U.S. Geological Survey, Mineral Commodity Summaries, January 2024. Accessed via website: https://pubs.usgs.gov/periodicals/mcs2024/mcs2024-cobalt.pdf on February 13, 2025.

WEF, 2024, "AI and Energy: Will AI Help Reduce Emissions or Increase Demand? Here's What to Know," World Economic Forum, July 22. Accessed via website: https://www.weforum.org/stories/2024/07/generative-ai-energy-emissions/ on February 13, 2025.

Wu, C.-J., Acun, B., Raghavendra, R. and K. Hazelwood, 2024, "Beyond Efficiency: Scaling AI Sustainably," arXiv.org, June 22. Accessed via website: https://arxiv.org/abs/2406.05303 on February 13, 2025.

Wu, J. and D. Bosa, 2025, "How China's New AI Model DeepSeek is Threatening U.S. Dominance," CNBC, January 24. Accessed via website: https://www.nbcnewyork.com/news/business/money-report/how-chinas-new-ai-model-deepseek-is-threatening-u-s-dominance/6120598/ on January 24, 2025.

Author Biographies

JODIE GUNZBERG, CFA
Managing Partner, InFi Strategies, U.S.A.

Jodie Gunzberg, CFA, is the Managing Partner of InFi Strategies, leading the firm in providing innovative indexing and investment strategies. With over 25 years of experience in the investment industry, she is a distinguished leader with buy-side, indexing, and investment consulting expertise.

Recently, Gunzberg was the Head of Strategic Partnerships at Cboe Global Markets, where she developed strategic partnerships and managed key commercial relationships to drive growth. She crafted policies and strategies for multi-party arrangements, expanded into new business lines, and established global fund issuer services. She was also on the Investment Committee overseeing Cboe's employee benefit plans.

Before Cboe, Jodie Gunzberg was a Managing Director at CoinDesk Indices, directing index product management and strategy, and serving as the company's public spokesperson. She defined digital asset industries, shaped index offerings, and created groundbreaking products for 401(k) plans and financial advisors.

Previously, Gunzberg was the Chief Institutional Investment Strategist at Morgan Stanley, leading the Graystone Investment Office. She developed institutional model portfolios and provided thought leadership for ultra-high-net-worth individuals, pensions, endowments, and foundations, managing assets totaling approximately $400 billion.

Earlier in her career, she was a Managing Director and Head of U.S. Equities at S&P Dow Jones Indices, overseeing the S&P 500 and Dow indices. Her analysis on women in the workforce's impact on global stock markets gained notable recognition. She was the most frequently quoted S&P Global spokesperson, and authored a blog recognized by *U.S. News & World Reports* as a top financial blog. She began at S&P DJI as the Director of Commodities Product Management.

Gunzberg is a CFA charterholder and a member of the CFA Institute's Education Advisory Committee. She previously served on the Boards of CFA NY and CFA Chicago, and founded the CFA New Jersey Group of CFA NY. Currently, she is on the Advisory Board of Hofstra University's Department of Finance. She is also an Editorial Board Member of the *Commodity Insights Digest*, a publication of Bayes Business School, City St George's, University of London, U.K., and she earned her M.B.A. from the University of Chicago Booth School of Business and a B.S. in Mathematics from Emory University.

HILARY TILL
Managing Co-Editor, *Commodity Insights Digest*; Honorary Senior Visiting Fellow, Bayes Business School, City St George's, University of London, U.K.; and Principal, Chicago-based Premia Research LLC, U.S.A.

Hilary Till began her career in the AI Group of what is now Accenture where she developed an expert system for Ford Motor Company. Relatedly, she is currently a member of the Bretton Woods Committee's AI Working Group.

Relevant to this article, Till has additionally lectured to members of the Chinese futures industry on the lessons of American economic history, given the early developmental experiences in the city of Chicago and the state of Illinois. She provided these lectures on behalf of the Chicago Institute of Investment.

Hilary Till has a B.A. with General Honors in Statistics with Concentrations in Computer Science and Economics from the University of Chicago where she studied under a full four-year tuition scholarship. In addition, she has an M.Sc. in Statistics from the London School of Economics where she studied under a private fellowship administered by the Fulbright Commission.

Reflecting her quantitative expertise, Till serves as a board member of New York-based International Association for Quantitative Finance and is a fellow of the Royal Statistical Society (U.K.). She can be followed on X at https://x.com/HTill42884.

COMMODITY INSIGHTS
DIGEST

Commodity Insights Digest (*CID*, ISSN 2996-654X) is a publication of Bayes Business School, City St George's, University of London (U.K.), in association with Chicago-based Premia Research LLC (U.S.A.). The digest is co-edited by Ana-Maria Fuertes, Ph.D. in International Finance, and Hilary Till, M.Sc. in Statistics.

CID seeks to foster knowledge transfer between academics and practitioners by publishing:

- Scholarly research digest articles;
- Practitioner insights and interviews;
- Articles on economic history; and
- Book reviews.

Contributors to *CID* include (a) scholars seeking to enhance the impact of their research by delving into the practical implications of their theoretical or empirical studies; (b) consultants reporting on challenges faced by commodity market participants; and (c) industry economists providing analyses on their areas of market expertise.

Bayes Business School, City St George's, University of London is a world-leading provider of education and cutting-edge research. Bayes' unique location enables strong links with industry in the City of London.

© 2025 Bayes Business School, City St George's, University of London (U.K.)

The cover image is excerpted from the announcement of a London event on "AI/ML in Finance," which was jointly organized by the International Association for Quantitative Finance and the London School of Economics (LSE) Data Science Institute. The event was held at the LSE's Shaw Library, and it took place on June 24, 2025. ■

BAYES BUSINESS SCHOOL
CITY ST GEORGE'S
UNIVERSITY OF LONDON

Physical Address
Bayes Business School
106 Bunhill Row London,
EC1Y 8TZ
United Kingdom

Website
bayes-cid.com

Contact
To submit articles, authors can contact the editors via editors@bayes-cid.com

City St George's is a member institution of the University of London

UNIVERSITY OF LONDON