Whispers in the oil market: exploring sentiment and uncertainty insights

Luigi Gifuni

University of Strathclyde

EIA workshop

March 19, 2025

Introduction

- **Accurate prediction** of **oil prices** is not merely market speculation but a **crucial** necessity for central banks to make **informed policy decisions** in a constantly volatile economic landscape.
- **Related literature:** Alquist and Kilian (2010); Baumeister and Kilian (2012, 2014, 2015); Baumeister et al. (2015).
- Oil price **forecast** has become notably **challenging after 2010** Baumeister et al. (2020).

Challenges

Current methodologies are **vulnerable** to **unforeseen events** impacting the oil market.

One **reason:** Central bank (and academic) researchers **forecast** real oil prices **using** official macroeconomic **hard data**, such as proxies of global real economy, oil production and inventories.

- $\Rightarrow$ They are made **available with delay**, and often **require** frequent **revisions** that prevent timely forecasts Baumeister and Kilian (2012); Alquist et al. (2013).
- $\Rightarrow$ They are by nature **slow to react** to economic changes.

Questions in the Paper

Can **text data** be leveraged to:

- provide **timely insights** and enable real-time forecasting?
- **predict** periods of **oil market instability**?
- **enhance** the accuracy of **monthly forecasts** for real oil prices?

Main Findings

- Oil market **sentiment indicators react more promptly** to events affecting oil prices than do oil price uncertainty measures.
- A **novel** text-based oil market sentiment **indicator** (**TOSI**) is introduced.
- **TOSI** significantly **enhances** the **forecasting accuracy** of real oil prices, particularly during periods of oil market instability.
- The findings remain **robust** across both **density forecast** evaluations and assessments of **directional accuracy**.

Text Data

- **6 million** daily **news** items from *The Financial Times*, *Thompson-Reuters*, and *The Independent*.
- Articles are downloaded from the LexisNexis database.
- Selection is based on the **co-occurrence** of the words **“oil”** and **“price”**.
- Articles are filtered by industry: *Banking & Finance, Energy & Utilities, Agriculture, Mining & Extraction, Market Research & Analysis, Manufacturing, and Retail*.
- Sample period: **January 1982 – June 2021**.
- Data cleaning follows Gentzkow et al. (2019).

Textual Analysis

- **Unigram count:**
  - ⇒ Computes word probabilities via maximum likelihood.
  - ⇒ "econom" for deriving sentiment and "uncertain" for uncertainty.
- **Dictionary methods:**
  - ⇒ Constructs sentiment indicators using pre-scored vocabularies.
  - ⇒ Employed dictionaries include Financial Stability Correa et al. (2017), Financial Liability Loughran and McDonald (2011), Afinn Nielsen (2011), and Harvard-IV.
- **Geometrical models:**
  - *TF:* Measures token frequency normalised by the total number of documents.
  - *TF-IDF:* Adjusts TFM to down-weight common terms.
- **Computer science:**
  - **VADER:** Rule-based similar to dictionary.
  - **Opinion Sentiment:** Employs a support vector machine classifier to categorise sentiment.
  - **BERT:** Combines pre-trained rule-based weights with a transformer architecture for contextual understanding.

Econometric method

Carriero et al. (2019)'s SV-BVAR.

Reduced form:

$$Y_t = c + \Phi_1 Y_{t-1} + ... + \Phi_p Y_{t-p} + \varepsilon_t, \qquad \varepsilon_t \sim N(0, \Sigma_t) \tag{1}$$

Structural version:

$$Y_t = c + \Phi_1 Y_{t-1} + ... + \Phi_p Y_{t-p} + A^{-1}\Lambda_t^{1/2}\epsilon_t, \qquad \epsilon_t \sim iid\ N(0, I_N) \tag{2}$$

$$\text{with } \Lambda_t = \begin{pmatrix} \exp(\lambda_{1,t}) & 0 & \cdots & 0 \\ 0 & \exp(\lambda_{2,t}) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \exp(\lambda_{N,t}) \end{pmatrix}, \tag{3}$$

where:

$$\Sigma_t = \left(A^{-1}\right)\Lambda_t\left(A^{-1}\right)' \qquad \text{and} \qquad \Sigma_t^{1/2} = \left(A^{-1}\right)\Lambda_t^{1/2}$$

$\lambda_{1,t}, \lambda_{2,t}, \ldots, \lambda_{N,t}$ refer to the log volatility of the N structural shocks in equation (1), and its law of motion is

$$\tilde{\lambda}_t = \tilde{\lambda}_{t-1} + \nu_t,$$

One text source

Table 1: Recursive MSPE ratios for real oil price forecasts: SVBVAR(12) models with alternative text-based indicators relative to a random walk.

Text index	h = 1 WTI	h = 1 BRENT	h = 3 WTI	h = 3 BRENT	h = 6 WTI	h = 6 BRENT	h = 12 WTI	h = 12 BRENT	h = 24 WTI	h = 24 BRENT
No - text	0.881**	0.888**	0.909	0.917	0.935	0.937	0.994	0.982	0.930	0.919
Sentiment count	0.883**	0.897**	0.925	0.942	0.943	0.975	1.011	1.013	0.938	0.932
Uncertainty count	0.896**	0.911*	0.926	0.943	0.963	0.969	1.019	1.041	0.936	**0.917**
Boolean count	0.888**	0.905*	0.917	0.932	0.948	0.943	0.995	0.997	0.939	0.923
Financial Stability	**0.879****	0.890**	0.935	0.961	0.993	1.017	1.041	1.053	0.991	0.978
Financial Liability	**0.871****	**0.884****	0.912	0.940	0.935	0.962	**0.959**	**0.977**	0.950	0.953
Afinn	**0.860****	**0.880****	**0.898**	0.920	**0.927**	**0.932**	**0.978**	0.988	**0.923**	0.939
Harvard-IV	**0.878****	0.891*	**0.907**	0.936	0.944	0.962	1.006	1.048	0.940	0.934
Sentiment tfm	**0.877****	**0.876****	**0.896**	**0.904**	**0.923**	**0.920**	**0.974**	**0.977**	0.948	0.924
Uncertainty tfm	**0.880****	0.895*	0.911	0.927	0.942	0.937	**0.989**	1.022	0.943	0.935
Sentiment tf-idf	**0.875****	**0.886****	0.910	0.926	0.955	0.963	1.047	1.034	1.016	0.983
Uncertainty tf-idf	**0.878****	**0.885****	0.922	0.921	0.939	0.947	**0.982**	1.010	0.959	0.948
VADER	**0.864****	**0.883****	**0.895**	**0.914**	0.939	0.938	1.019	1.034	**0.914**	**0.902**
Opinion Sentiment	**0.870****	**0.882****	0.913	**0.909**	**0.929**	**0.915**	**0.962**	0.986	0.949	**0.913**
BERT	**0.834*****	**0.857****	**0.870**	**0.890**	**0.901**	**0.901**	**0.941**	**0.925**	0.940	0.925

Note: In column 1, the No-text row corresponds to the baseline Baumeister et al. (2020)'s SVBVAR model without text data. Bold values report MSPEs of text-based SVBVARs outperforming MSPEs of no-text-based SVBVAR for each specific time horizon *h* and oil price measure (WTI and Brent). *, **, and *** denote significance levels of 10%, 5%, and 1%, respectively, when tested against the random walk model, while blue, red, and yellow cells denote significance levels of 10%, 5%, and 1%, respectively, when tested against the No-text-based SVBVAR model. Significance tests are performed using the Diebold–Mariano test.

Multiple text sources

Table 2: Recursive MSPE ratios relative to a random walk forecast of alternative monthly indicators of real oil prices. Combination between different text data sources

Text sources	**h = 1** WTI	BRENT	**h = 3** WTI	BRENT	**h = 6** WTI	BRENT	**h = 12** WTI	BRENT	**h = 24** WTI	BRENT
					VADER					
FT-TR	0.858**	0.885**	0.886	0.916	0.915	0.940	0.983	1.002	0.898	0.884
FT-IND	0.865**	0.883**	0.901	0.925	0.939	0.938	1.004	1.007	0.912	0.897
FT-IND-TR	0.865**	0.883**	0.894	0.928	0.931	0.941	0.988	0.985	0.899	0.885
					Opinion Sentiment					
FT-TR	0.870**	0.881**	0.908	0.925	0.933	0.917	0.981	0.958	0.938	0.908
FT-IND	0.875**	0.891*	0.895	0.906	0.906	0.893	0.923	0.940	0.918	0.902
FT-IND-TR	0.869**	0.884**	0.893	0.903	0.905	0.884	0.946	0.941	0.951	0.939
					BERT					
FT-TR	0.829***	0.831***	0.858	0.861	0.897	0.881	0.935	0.922	0.926	0.911
FT-IND	0.848**	0.867**	0.878	0.895	0.902	0.898	0.944	0.921	0.934	0.915
FT-IND-TR	0.853**	0.865**	0.878	0.887	0.909	0.899	0.961	0.913	0.943	0.930

Note: In column 1, FT: Financial Times, TR: Thomson Reuters, IND: Independent. Blue values report the lowest MSPE results relative to a specific time horizon *h* and oil price measure (WTI and Brent). *, **, *** respectively denote 10%, 5% and 1% level of significance as suggested by the Diebold-Mariano test.

MSPE common component

Table 3: Recursive MSPE of alternative monthly indicators of real oil prices in a SVBVAR(12). Comparison between different predictors.

	h = 1		h = 3		h = 6		h = 12		h = 24	
Predictor	**WTI**	**Brent**	**WTI**	**Brent**	**WTI**	**Brent**	**WTI**	**Brent**	**WTI**	**Brent**
PCA(1)	**0.838*****	**0.829*****	**0.858**	0.884	**0.899**	**0.897**	**0.954**	**0.937**	**0.916**	**0.900**
FVAR	0.840**	0.859**	0.870	**0.881**	0.910	0.901	0.956	0.943	0.929	0.909
3PRF	0.869**	0.871**	0.888	0.890	0.913	0.902	1.005	0.986	0.950	0.906
OPU	0.887**	0.881**	0.929	0.895	0.958	0.928	0.970	0.959	0.919	0.904
BW	0.884**	0.898**	0.918	0.928	0.955	0.945	1.009	0.993	0.977	0.988
IV – *VADER*	0.879**	0.888*	0.890	0.925	0.911	0.932	0.961	0.988	1.043	1.044
IV – *Opinion Sentiment*	0.879*	0.892*	0.895	0.915	0.915	0.934	0.966	0.989	1.049	1.054
IV – *Bert*	0.877**	0.888*	0.891	0.913	0.910	0.933	0.964	0.990	1.038	1.047

Note: Bold values show the lowest MSPE results relative to a specific time horizon *h* and oil price measure (WTI and Brent). *, **, and *** denote significance levels of 10%, 5%, and 1%, respectively, when tested against the random walk model, while blue, red, and yellow cells denote significance levels of 10%, 5%, and 1%, respectively, when tested against the No-text-based SVBVAR model. Significance tests are performed using the Diebold–Mariano test.

MCS common component

Table 4: Recursive MCS of alternative monthly indicators of real oil prices in a SVBVAR(12). Comparison between different predictors.

	h = 1		h = 3		h = 6		h = 12		h = 24	
Predictor	**WTI**	**Brent**	**WTI**	**Brent**	**WTI**	**Brent**	**WTI**	**Brent**	**WTI**	**Brent**
PCA(1)	**1.000**	**1.000***	**1.000**	0.993	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**
FVAR	0.961	0.000	0.360	**1.000**	0.281	0.999	0.996	0.975	0.397	0.735
3PRF	0.697	0.363	0.985	1.000	1.000	1.000	0.992	0.980	0.995	1.000
OPU	0.255	0.158	0.460	0.996	0.578	0.917	0.996	0.990	1.000	1.000
BW	0.675	0.167	0.657	0.703	0.818	0.910	0.983	0.940	0.983	0.568
IV – *VADER*	0.081	0.082	0.404	0.032	1.000	0.851	1.000	0.918	0.012	0.144
IV – *Opinion Sentiment*	0.064	0.075	0.206	0.181	0.838	0.822	0.980	0.905	0.008	0.145
IV – *Bert*	0.088	0.129	0.316	0.234	1.000	0.851	0.999	0.894	0.012	0.190

Note: Bold values highlight models included in the Superior Set of Models, as identified by the MCS procedure at the 1% confidence level. *, **, *** respectively denote 10%, 5% and 1% level of significance, for a test of the null hypothesis that there is no difference in predictive accuracy among the models.

TOSI

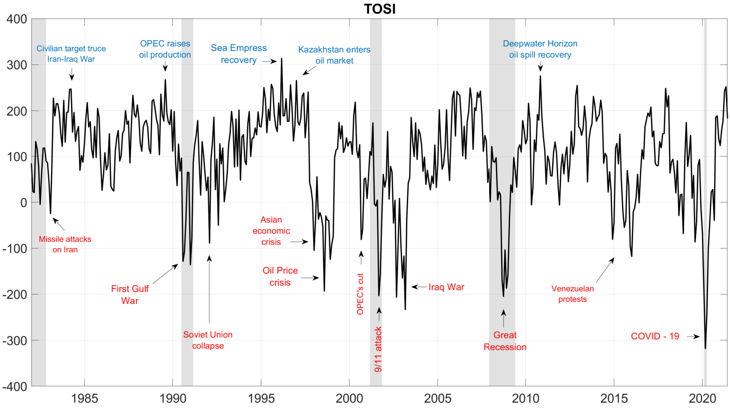


Figure 1: Common component extracted from a dataset comprising VADER and BERT text indicators. Blue and red notes describe historical events that positively or negatively impacted oil prices. Shaded areas represent recessions as identified by the NBER. The sample period covers January 1982 to June 2021, with the index normalized to a mean of 100.

Cumulative sum of forecasting errors (CSFE)

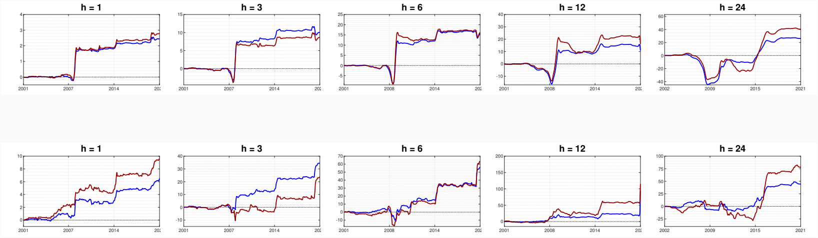


Figure 2: CSFE difference: top row TOSI-based model versus a random walk. Bottom row TOSI-based model versus model without text data. Blue and red lines describe WTI and Brent crude oil prices respectively.

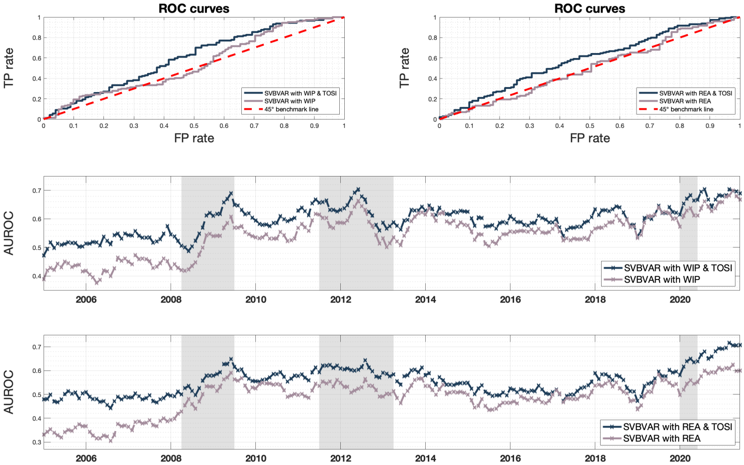
Density forecast

Table 5: Recursive ALPL (Panel A) and ACRPS (Panel B) of alternative monthly indicators of real oil prices across various SVBVAR(12) models.

Panel A: ALPL		h = 1		h = 3		h = 6		h = 12		h = 24	
Variables	**Model**	**WTI**	**Brent**	**WTI**	**Brent**	**WTI**	**Brent**	**WTI**	**Brent**	**WTI**	**Brent**
p^{oil}, *TOSI*	2-SVBVAR(12)	1.021***	**0.143***	0.206***	-0.102***	-0.168***	-0.377***	**-0.482***	-0.643***	**-0.777***	**-0.945***
GCOP, p^{oil}, *TOSI*	3-SVBVAR(12)	1.016***	0.122***	0.202***	-0.124***	-0.169***	-0.397***	-0.488***	-0.653***	-0.784***	-0.956***
GCOP, *WIP*, p^{oil}, *TOSI*	4-SVBVAR(12)	1.018***	0.084***	0.226***	-0.146***	-0.159***	-0.418***	-0.508***	-0.692***	-0.815***	-0.987***
GCOP, *WIP*, *Oinv*, p^{oil}, *TOSI*	5-SVBVAR(12)	**1.030***	0.099***	**0.245***	**-0.075***	**-0.144***	**-0.341***	-0.492***	**-0.634***	-0.812***	-0.973***
Panel B: ACRPS		h = 1		h = 3		h = 6		h = 12		h = 24	
Variables	**Model**	**WTI**	**Brent**	**WTI**	**Brent**	**WTI**	**Brent**	**WTI**	**Brent**	**WTI**	**Brent**
p^{oil}, *TOSI*	2-SVBVAR(12)	0.062***	0.114	0.129***	0.171	0.182***	0.226	**0.247***	**0.296**	**0.307***	0.374
GCOP, p^{oil}, *TOSI*	3-SVBVAR(12)	0.062***	0.115	0.130***	0.174	0.183***	0.230	0.248***	0.301	0.312	0.384
GCOP, *WIP*, p^{oil}, *TOSI*	4-SVBVAR(12)	0.061***	0.116	0.128***	0.176	0.182***	0.233	0.256	0.311	0.328	0.398
GCOP, *WIP*, *Oinv*, p^{oil}, *TOSI*	5-SVBVAR(12)	**0.060***	**0.112**	**0.125***	**0.167**	**0.176***	**0.219**	0.248***	**0.296**	0.308*	**0.371**

Note: In column 1: p^{oil} = oil prices, *GCOP* = global crude oil production, *WIP* = world industrial production, *Oinv* = oil inventories, *TOSI* = text oil sentiment indicator. For negative results, a less negative value indicates that the model offers a more accurate representation of the true density forecast. Bold values indicate the best model relative to a specific time horizon h and oil price measure (WTI and Brent). *, **, and *** denote significance levels of 10%, 5%, and 1%, respectively, when tested against the random walk model, while blue, red, and yellow cells denote significance levels of 10%, 5%, and 1%, respectively, when tested against the equivalent model without TOSI. Significance tests are performed using the Diebold–Mariano test.

Directional accuracy 1-month ahead forecast



Directional accuracy over multiple subsamples

Table 6: AUROC estimates.

Economic Period	Time Frame	4-SVBVAR	TOSI-based 5-SVBVAR
Global financial crisis	2008M4 – 2009M7	0.7857**	**0.8036*****
Eurozone sovereign debt crisis	2011M7 – 2013M4	0.4636	**0.5909**
COVID-19 recession	2020M1 – 2020M6	0.6250	**0.8750****
Full sample	1982M1 – 2021M6	0.5436	**0.6044*****

Note: Recession dates are determined by the CEPR-EABCN Euro Area Business Cycle Dating Committee. Bold values indicate superior predictive performance. *, **, and *** respectively denote 10%, 5%, and 1% level of significance, for a test of the null hypothesis that oil price forecasts are indistinguishable from a pure noise signal.

Sensitivity analysis

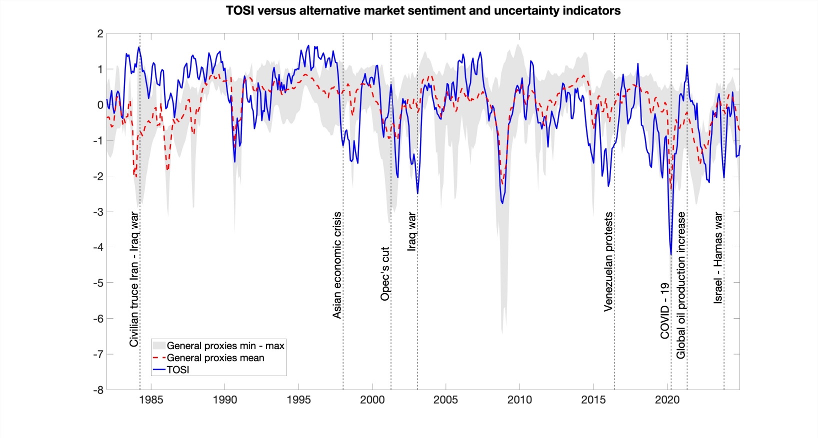


Figure 3: Three-month moving average of TOSI is plotted alongside the three-month moving average of five established sentiment and uncertainty proxies. The shaded band indicates the range (minimum to maximum) of these proxies at each time point, illustrating that while the general sentiment measures track overall market conditions, TOSI exhibits sharper, oil-specific reactions during key market events.

Sensitivity analysis

Table 7: Recursive MSPE for one-month-ahead forecasts comparing TOSI-based models to those using alternative market sentiment and uncertainty indicators during recent oil market events (WTI crude oil prices).

Comparison for WTI crude oil prices	**Covid-19** Feb-2020–Dec-2021	**Russia-Ukraine War** Feb-2022–Jul-2023	**Israel-Hamas War** Aug-2023–Dec-2023
TOSI vs. OVXCLS	0.868*	0.727	0.396***
TOSI vs. Baker and Wurgler (2006)'s IS	0.883***	0.972	0.709
TOSI vs. Baker et al. (2016)'s Global EPU	0.834***	0.824	0.438**
TOSI vs. Bekaert et al. (2022)'s RA	0.997	0.757**	0.690**
TOSI vs. Abiad and Qureshi (2023)'s OPU	0.925**	0.860	0.636**

Note: General indicators serve as the benchmark models. IS = investor sentiment; EPU = economic policy uncertainty; RA = risk aversion; OPU = oil price uncertainty. The reported values are relative MSPEs, computed as the ratio of the MSPE of the TOSI-based model to that of the benchmark model. Values below one indicate that the TOSI-SVBVAR model yields more accurate forecasts of WTI crude oil prices. *, **, *** respectively denote 10%, 5% and 1% level of significance as suggested by the Diebold-Mariano test.

Sensitivity analysis

Table 8: Recursive MSPE for one-month-ahead forecasts comparing TOSI-based models to those using alternative market sentiment and uncertainty indicators during recent oil market events (Brent crude oil prices).

Comparison for Brent crude oil prices	**Covid-19** Feb-2020–Dec-2021	**Russia-Ukraine War** Feb-2022–Jul-2023	**Israel-Hamas War** Aug-2023–Dec-2023
TOSI vs. OVXCLS	0.892	0.725**	0.355***
TOSI vs. Baker and Wurgler (2006)'s IS	0.892**	0.958	0.553**
TOSI vs. Baker et al. (2016)'s Global EPU	0.849**	0.789**	0.389*
TOSI vs. Bekaert et al. (2022)'s RA	0.996	0.740***	0.690*
TOSI vs. Abiad and Qureshi (2023)'s OPU	0.934	0.845**	0.523***

Note: General indicators serve as the benchmark models. IS = investor sentiment; EPU = economic policy uncertainty; RA = risk aversion; OPU = oil price uncertainty. The reported values are relative MSPEs, computed as the ratio of the MSPE of the TOSI-based model to that of the benchmark model. Values below one indicate that the TOSI-SVBVAR model yields more accurate forecasts of Brent crude oil prices *, **, *** respectively denote 10%, 5% and 1% level of significance as suggested by the Diebold-Mariano test.

Key Takeaways

- Text data on oil news presents both strengths and weaknesses.
- **Weaknesses**: Uncertainty measures do not provide reliable forecasts of real oil prices.
- **Strengths**: Sentiment metrics show strong potential for improving the forecasting accuracy of real oil prices.
- **Combination**: Using multiple text sources does not necessarily lead to better information.
- **TOSI**: Can enhance the accuracy of real oil price forecasts, particularly during periods of oil market instability.
- **Latest TOSI Release**: Available at https://sites.google.com/view/luigigifuni/research.

References

Abiad, A. and Qureshi, I. A. (2023). The macroeconomic effects of oil price uncertainty. *Energy Economics*, 125:106839.

Alquist, R. and Kilian, L. (2010). What do we learn from the price of crude oil futures? *Journal of Applied Econometrics*, 25(4):539–573.

Alquist, R., Kilian, L., and Vigfusson, R. J. (2013). Forecasting the price of oil. In *Handbook of Economic Forecasting*, volume 2, pages 427–507. Elsevier.

Baker, M. and Wurgler, J. (2006). Investor sentiment and the cross-section of stock returns. *The Journal of Finance*, 61(4):1645–1680.

Baker, S. R., Bloom, N., and Davis, S. J. (2016). Measuring economic policy uncertainty. *The Quarterly Journal of Economics*, 131(4):1593–1636.

Baumeister, C., Guérin, P., and Kilian, L. (2015). Do high-frequency financial data help forecast oil prices? The MIDAS touch at work. *International Journal of Forecasting*, 31(2):238–252.

Baumeister, C. and Kilian, L. (2012). Real-time forecasts of the real price of oil. *Journal of Business & Economic Statistics*, 30(2):326–336.

Baumeister, C. and Kilian, L. (2014). What central bankers need to know about forecasting oil prices. *International Economic Review*, 55(3):869–889.

Baumeister, C. and Kilian, L. (2015). Forecasting the real price of oil in a changing world: a forecast combination approach. *Journal of Business & Economic Statistics*, 33(3):338–351.

Baumeister, C., Korobilis, D., and Lee, T. K. (2020). Energy markets and global economic conditions. *Review of Economics and Statistics*, pages 1–45.

Bekaert, G., Engstrom, E. C., and Xu, N. R. (2022). The time variation in risk appetite and uncertainty. *Management Science*, 68(6):3975–4004.

Carriero, A., Clark, T. E., and Marcellino, M. (2019). Large Bayesian vector autoregressions with stochastic volatility and non-conjugate priors. *Journal of Econometrics*, 212(1):137–154.

Correa, R., Garud, K., Londono, J. M., Mislang, N., et al. (2017). Constructing a dictionary for financial stability. *IFDP notes. Board of Governors of the Federal Reserve System, Washington, DC*.

Gentzkow, M., Kelly, B., and Taddy, M. (2019). Text as data. *Journal of Economic Literature*, 57(3):535–574.

Loughran, T. and McDonald, B. (2011). When is a liability not a liability? textual analysis, dictionaries, and 10-ks. *The Journal of finance*, 66(1):35–65.

Nielsen, F. Å. (2011). A new anew: Evaluation of a word list for sentiment analysis in microblogs. *arXiv preprint arXiv:1103.2903*.