

Forecasting Oil-Price Tail Risk with Newspaper Attention

Gabriel Appau Abeyie^a

^aDepartment of Economics and Finance, University of North Carolina Wilmington, 601 S. College Rd., Wilmington, NC, 28403, USA

ARTICLE INFO

Keywords:

Oil price volatility
Newspaper attention
Text analysis
Volatility forecasting
Value-at-Risk
Downside risk

ABSTRACT

This paper examines oil-news attention as a predictor of crude oil price volatility. Using oil-related news articles from four major U.S. newspapers between 1988 and 2024, we construct monthly measures of oil-news attention and evaluate their out-of-sample forecasting power for WTI and Brent realized volatility. Adding oil-news attention to autoregressive (AR) and heterogeneous autoregressive (HAR) benchmarks significantly improves forecast accuracy, particularly for downside volatility. The improvements persist even when the benchmark already includes implied volatility from the VIX or the OVX, and they extend to 5% Value-at-Risk (VaR) forecasts. Consistent with the mixture-of-distributions hypothesis, what matters for forecasting oil-price risk is how much the press writes about oil, not what it says.

1. Introduction

Crude oil is an important global commodity, and higher oil prices can fuel inflation across economies. The ability to forecast oil prices is therefore valuable for policymakers and businesses. However, oil prices are influenced by a wide range of international and macroeconomic factors, which makes them notoriously difficult to forecast. At monthly horizons, oil returns are close to unpredictable, and forecasts of the price level often struggle to outperform a no-change benchmark (Hamilton, 2009; Kilian, 2009; Baumeister and Hamilton, 2019; Alquist, Kilian and Vigfusson, 2013; Baumeister and Kilian, 2015). Oil-price volatility, however, is more forecastable, and it plays a central role in risk management, hedging, and the pricing of oil-linked claims. A large literature has developed models for forecasting volatility, including applications to oil markets (Andersen, Bollerslev, Diebold and Labys, 2003; Corsi, 2009; Sévi, 2014; Bollerslev, Li and Zhao, 2020).

Most of these models rely on market-based and macroeconomic predictors. Text data offer an alternative, and their use in forecasting has grown in recent years, with several studies finding useful predictive information in them (Lima and Godeiro, 2023; Barbaglia, Consoli and Manzan, 2023; Hong, Jiang, Meng and Xue, 2025; Wei, Lima and Godeiro, 2026). In the oil market, recent work has examined the predictive power of the content of oil news (Li, Shang and Wang, 2019; Li, Jiang, Li and Wang, 2021; Abeyie, 2026). The attention that newspapers devote to oil, however, has not been studied as a predictor of oil-price volatility. Oil-market coverage is a natural candidate: when oil markets are under stress, oil news crowds the business pages, whereas in calm periods it recedes. Because the amount of coverage moves with oil-market stress, it may carry information about future volatility.

In this paper, we ask whether the volume of oil-market coverage in major U.S. newspapers improves out-of-sample

forecasts of oil-price volatility relative to conventional benchmarks.¹ We also ask whether any improvement is concentrated in downside risk, which we measure using downside realized semivariance, the part of realized variance constructed from negative daily returns.

To do this, we collect over 79,000 oil-related news articles from four prominent U.S. newspapers, the *Wall Street Journal*, *New York Times*, *Washington Post*, and *Los Angeles Times*, over the period 1988–2024. We preprocess the text and construct four monthly oil-news indices: an oil-news attention index, which measures the volume of oil coverage; an oil-risk attention index, which counts risk and disruption language near oil-market references; an oil-uncertainty index, which measures the frequency of uncertainty language in text surrounding oil-market references; and an oil-specific price-pressure tone index, which measures the balance between language pointing to upward and downward pressure on oil prices.

The predictive performance of these indices is evaluated in a recursive out-of-sample framework with an expanding estimation window, in which we forecast one-month-ahead realized volatility and downside realized volatility for both WTI and Brent.² The benchmarks are an autoregressive (AR) realized-volatility model and the heterogeneous autoregressive (HAR) model of Corsi (2009). The HAR model predicts next month's realized volatility from realized volatility in the current month and its averages over the past three and twelve months, capturing short-, medium-, and long-run volatility dynamics. Forecast accuracy is measured by the out-of-sample R^2 relative to the corresponding benchmark, and significance is assessed with the Clark and West (2007) test for nested models.

Our main finding is that oil-news attention improves out-of-sample forecasts of oil-price volatility, with the gains concentrated in downside risk. Adding oil-news attention to

¹Although crude oil trades in a global market, the national U.S. newspapers we use cover that global market, including OPEC decisions, geopolitical developments, and shifts in global demand, rather than a purely domestic one.

²West Texas Intermediate (WTI) and Brent are the two main international benchmarks for crude oil prices.

the AR and HAR benchmarks raises their out-of-sample R^2 for downside realized volatility for both WTI and Brent, and the improvement is statistically significant. It survives after controlling for the same-month return and its absolute value, and after adding option-implied volatility to the benchmark through the VIX and the oil-specific OVX. Oil-news attention thus has predictive content that option-implied-volatility benchmarks do not capture. The gains also carry over to the lower tail of the return distribution: oil-news attention improves one-month-ahead Value-at-Risk forecasts constructed from the volatility forecasts and evaluated with a quantile loss.

Two further results discipline the interpretation of this signal. First, it is the volume of oil coverage, not its tone, that carries information. The oil-specific price-pressure tone, built to capture the direction in which news points for oil prices, does not forecast volatility, so the directional content of oil-news language is uninformative about how much prices will move. Second, once the volume of coverage is controlled for, the risk-weighted version of that volume adds little. Oil-news attention and risk attention are highly correlated (0.93), and although risk attention has slightly more forecasting power than plain attention against volatility-history and other risk-proxy benchmarks (Section 4), this advantage disappears once option-implied volatility is included. The robust predictor is thus the volume of oil coverage. For forecasting risk, what matters is the salience of oil news, not its sentiment.

This paper contributes to three literatures. First, it adds to the literature on oil-price volatility forecasting (Sévi, 2014; Bollerslev et al., 2020) by introducing newspaper coverage volume as a predictor, showing that its predictive content is concentrated in downside and tail risk, and evaluating the forecasts with quantile and Expected-Shortfall loss functions suited to tail risk. The volume of coverage carries risk information beyond option-implied volatility and beyond standard measures of uncertainty and geopolitical risk (Baker, Bloom and Davis, 2016; Jurado, Ludvigson and Ng, 2015; Caldara and Iacoviello, 2022).

Second, the paper links to the literature on news sentiment and asset prices (Tetlock, 2007; Garcia, 2013) and on investor attention in asset markets (Da, Engelberg and Gao, 2011; Andrei and Hasler, 2015). We study whether the amount of oil coverage is informative about future risk. A key distinction is that our measure captures the supply of news, whereas search-based measures capture the demand for it: oil coverage reflects editorial decisions about events in the oil market. We therefore interpret the volume of coverage as a proxy for the intensity of information reaching the public, rather than as a direct measure of investor attention.

Finally, the paper contributes to work on measuring oil-related narratives and information from text (Shiller, 2017; Loughran, McDonald and Pragidis, 2019; Caldara and Iacoviello, 2022; Bybee, Kelly, Manela and Xiu, 2024). Following the context-dependent approach of Loughran et al. (2019), we build an oil-specific tone measure that distinguishes upward from downward price pressure. Horse-race

and encompassing tests then show how the information in directional tone and risk weighting compares with that in coverage volume.

The rest of the paper is organized as follows. Section 2 describes the text corpus, the text measures, and the realized-volatility data. Section 3 sets out the predictive regressions and the out-of-sample forecasting procedure. Section 4 reports the main results, including the decomposition of the signal into volume and risk content and the Value-at-Risk evaluation. Section 5 collects the robustness checks, and Section 6 concludes.

2. Data

This section describes the newspaper corpus, the four text-based measures (which we also refer to as indices), and the realized-volatility variables. The timing is strictly predictive: each measure for month t uses only articles published in month t and is used to forecast realized oil-price volatility in month $t + 1$.

2.1. Oil-News Corpus

The corpus is compiled from ProQuest TDM Studio and covers four major U.S. newspapers: the *Wall Street Journal*, the *New York Times*, the *Washington Post*, and the *Los Angeles Times*.³ The sample runs from January 1988 to December 2024. We keep all reported news articles, that is, news stories, feature stories, general stories, and lead stories, and exclude editorials and opinion pieces, since we aim to capture news reporting on oil markets rather than commentary on them.

The search targets the oil market rather than the broader energy sector. An article enters the raw corpus if it contains an oil-market term, such as “crude oil,” “oil price,” “OPEC,” “Brent crude,” or “WTI,” in combination with a market-relevant term such as “price,” “supply,” “demand,” “inventory,” “production,” “embargo,” “sanction,” “war,” “hurricane,” “refineries,” or “Strategic Petroleum Reserve.” The search returns 84,212 articles. After removing duplicate articles, articles without date or title information, and very short articles, we obtain 79,330 articles.

Table 1 reports the corpus composition by source. The *Wall Street Journal* contributes the largest share (44.5%), followed by the *New York Times* (24.2%) and the *Los Angeles Times* (21.6%); coverage of the *Washington Post* in the database begins in late 1996, leaving it with the smallest share. The two national business and general-news papers thus account for roughly 70% of the corpus.

Our baseline corpus is intentionally broad, because different oil-market events produce news of a very different nature: the Gulf War, the growth of the early 2000s, the 2008

³We use several newspapers because any single outlet’s oil coverage reflects idiosyncratic editorial choices, such as desk preferences, space constraints, and archival coverage, as well as the underlying flow of oil news; averaging across outlets keeps the common component and filters out outlet-specific noise, as in Baker et al. (2016) and Caldara and Iacoviello (2022).

spike and collapse, the 2014–2016 decline, the pandemic-related crash, and the recent Russia–Ukraine war are very different episodes. Narrowing the search would miss some of that diversity. We also construct a stricter corpus that drops recurring market-digest titles and requires either an oil-related headline or several oil mentions in the body. This corpus contains 35,455 articles and serves as a robustness sample.⁴

Although crude oil trades in a global market, we use national U.S. newspapers for three reasons. First, the major U.S. national press covers the global oil market rather than a domestic one: the *Wall Street Journal* and the *New York Times* report extensively on OPEC, Middle East geopolitics, Russian and Chinese supply and demand, and other international developments. Second, our measure is meant to capture the intensity of oil-related information reaching the public, and that information is largely global in content regardless of where it is printed. Third, and consistent with this, oil-news attention forecasts the volatility of Brent, the international benchmark, as well as that of WTI (Section 4); were the signal specific to U.S. conditions, it would not predict the global benchmark. We use these four outlets in particular because ProQuest TDM Studio provides long, consistent, full-text archives for them back to 1988, which is essential for a monthly panel spanning nearly four decades; Appendix A.4 shows that the signal is present across all four sources.

Table 1
Composition of the oil-news corpus by newspaper

Newspaper	Articles	Share (%)	Articles per month	Coverage
Wall Street Journal	35,278	44.5	79.5	1988–2024
New York Times	19,197	24.2	43.2	1988–2024
Los Angeles Times	17,165	21.6	40.1	1988–2024
Washington Post	7,690	9.7	22.8	1996–2024
Total	79,330	100.0	178.7	1988–2024

Notes: Composition of the cleaned baseline oil-news corpus (79,330 articles, January 1988–December 2024) by source. Articles are oil-query articles after cleaning and de-duplication. Articles per month is the mean over months in which the newspaper has at least one article; for the total, it is the mean over all 444 sample months. The oil-focused subset used to construct the attention measure contains 49,547 articles, with the same ordering of sources. Coverage of the *Washington Post* begins in late 1996. Source: ProQuest TDM Studio.

2.2. Textual Measures

From the corpus, we construct four monthly text indices that measure how much attention newspapers pay to the oil market and the character of that coverage. The first measures attention; the second, risk language; the third, uncertainty language; and the fourth, directional price-pressure tone. Attention is the main measure; the other three let us test whether these news characteristics add predictive power.

⁴Construction details are in Appendix A.1.

Oil-news volume is the number of articles that mention the oil market. The attention index is

$$\text{Attention}_t = \log(1 + N_t), \quad (1)$$

where N_t is the number of oil articles in month t .⁵

The risk-attention index measures how much risk- and uncertainty-related language appears near oil-market references. It uses the same dictionary as the uncertainty index, counting those words within symmetric ± 20 -word windows around each oil-market mention:

$$\text{RiskAttention}_t = \log(1 + \text{RiskHits}_t), \quad (2)$$

where RiskHits_t is the monthly count of dictionary words in those windows. Risk attention and uncertainty share this dictionary and differ only in normalization: risk attention is a log count, capturing the salience (the amount) of such language, whereas the uncertainty index divides the count by the number of window words, capturing its intensity per word. Risk attention can rise because oil coverage expands, because that coverage becomes more risk-focused, or both; Section 4.3 separates these channels.⁶

The uncertainty index uses the same dictionary, normalized by the number of words in the oil windows. With UncertaintyHits_t the count of dictionary words and WindowWords_t the number of words in the oil windows,

$$\text{Uncertainty}_t = \frac{\text{UncertaintyHits}_t}{\text{WindowWords}_t + 1}. \quad (3)$$

Tone is based on whether a term pushes oil prices up or down, regardless of its ordinary sentiment. Words like “cut,” “shortage,” “outage,” and “disruption” are negative-polarity words in a standard sentiment dictionary, but in oil-market articles they usually signal tighter supply and rising prices. We therefore classify terms into those indicating upward price pressure (production cut, supply disruption, shortage, sanction, war, hurricane, outage, strong demand) and those indicating downward price pressure (glut, oversupply, surplus, inventory build, production increase, weak demand, recession, economic slowdown). Terms whose polarity depends on context are classified using the oil-specific rules of Loughran et al. (2019). With UpPressure_t and DownPressure_t the monthly counts of upward- and downward-pressure terms and TotalWords_t the number of relevant terms, the tone index is

$$\text{Tone}_t = \frac{\text{UpPressure}_t - \text{DownPressure}_t}{\text{TotalWords}_t + 1}, \quad (4)$$

where positive and negative values indicate net upward and net downward price pressure in that month. Because Tone_t is a ratio, it is independent of the volume of oil-market news by construction.

⁵The logarithm dampens months of extreme news volume while preserving the information in spikes.

⁶The dictionary comprises risk- and uncertainty-related words such as risk, uncertainty, volatility, instability, concern, fear, doubt, speculation, tension, and threat; the full list is in Appendix A.1.

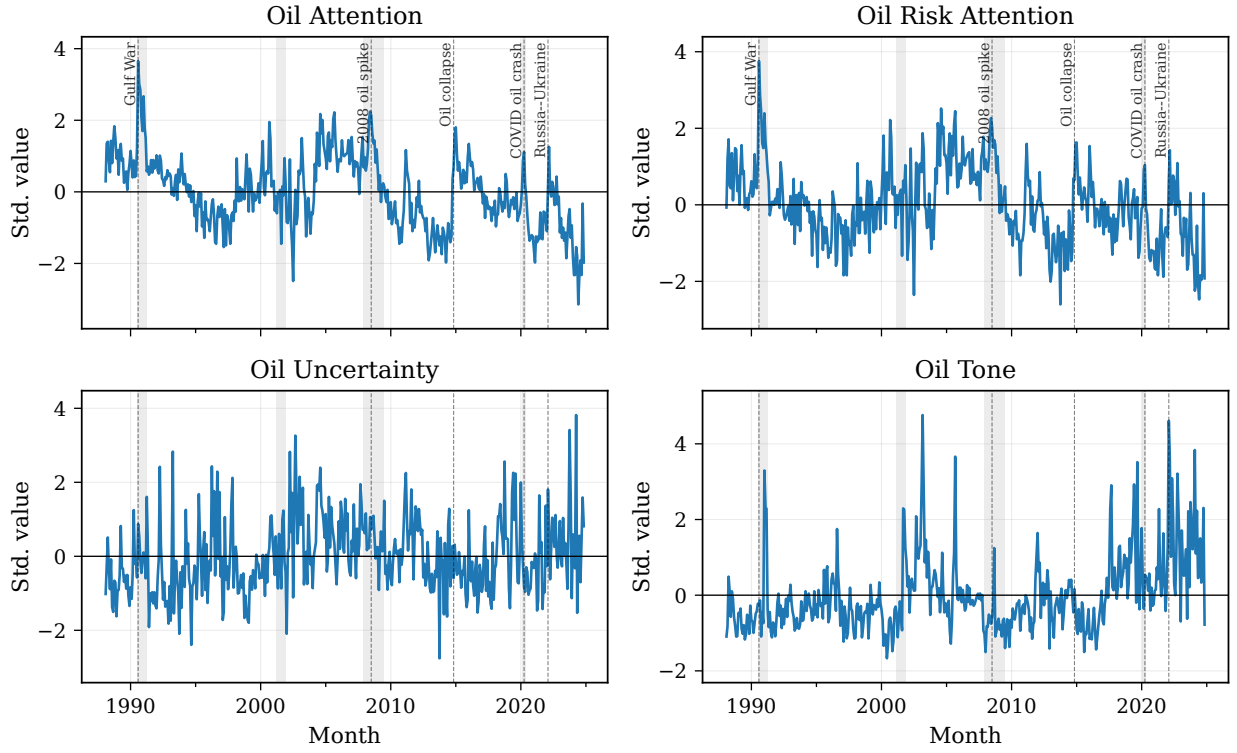


Figure 1: Oil narrative indices

Notes: Standardized monthly oil narrative indices, 1988:02–2024:11. Gray bands mark NBER recessions. Dashed vertical lines mark selected oil-market events: the Gulf War (1990), the Iraq War (2003), the 2008 price spike, the 2014–2016 price collapse, the COVID-19 shock (2020), and the Russia–Ukraine war (2022).

All indices are standardized to zero mean and unit variance before entering the regressions.⁷

2.3. Do the Measures Track Oil-Market Stress?

Figure 1 plots the four indices. Oil attention and risk attention rise around the episodes one would expect, including the Gulf War, the 2008 spike and collapse, the 2014–2016 decline, the COVID-19 shock, and the Russia–Ukraine war. This is descriptive evidence that the indices respond to oil-market stress; the forecasting tests follow.

2.4. Oil Prices and Realized Volatility

The price data are daily WTI and Brent crude prices. Daily log returns are

$$r_d = 100 \times (\log P_d - \log P_{d-1}), \quad (5)$$

and monthly realized variance is the sum of squared daily returns,

$$RV_t = \sum_{d \in t} r_d^2, \quad (6)$$

⁷This normalization introduces no look-ahead bias in the out-of-sample exercises. Because each forecasting regression includes an intercept, an affine transformation of a predictor does not change the forecast; we also confirm the results using recursive standardization that uses only information available up to each forecast origin.

with baseline outcome $\log(1 + RV_t)$. Downside realized variance uses only negative daily returns,

$$DRV_t = \sum_{d \in t} r_d^2 \mathbf{1}\{r_d < 0\}, \quad (7)$$

with outcome $\log(1 + DRV_t)$. Using both WTI and Brent lets us check that the predictive content is general across the two benchmarks.

The April 2020 WTI episode requires care, since negative prices make log returns undefined. We compute daily log returns only when the current and lagged prices are both strictly positive, and set the rest to missing. This affects only April 2020 for WTI; Brent stays positive throughout. The robustness analysis also excludes the COVID crash entirely.

Figure 2 plots the volatility outcomes. WTI and Brent volatility move together; downside volatility spikes sharply around the 2008 collapse, the COVID shock, and the Russia–Ukraine war; and all four series are persistent but episodic, which motivates the AR and HAR benchmarks.

Narratives in month t are matched to volatility in month $t + 1$, so the text is predetermined relative to the outcome. Because the baseline model already conditions on current volatility, the question is whether oil news adds information beyond RV_t . Table 2 reports descriptive statistics. The standardized indices are mean-zero with substantial variation, and downside RV has a lower mean than total RV but is no less variable.

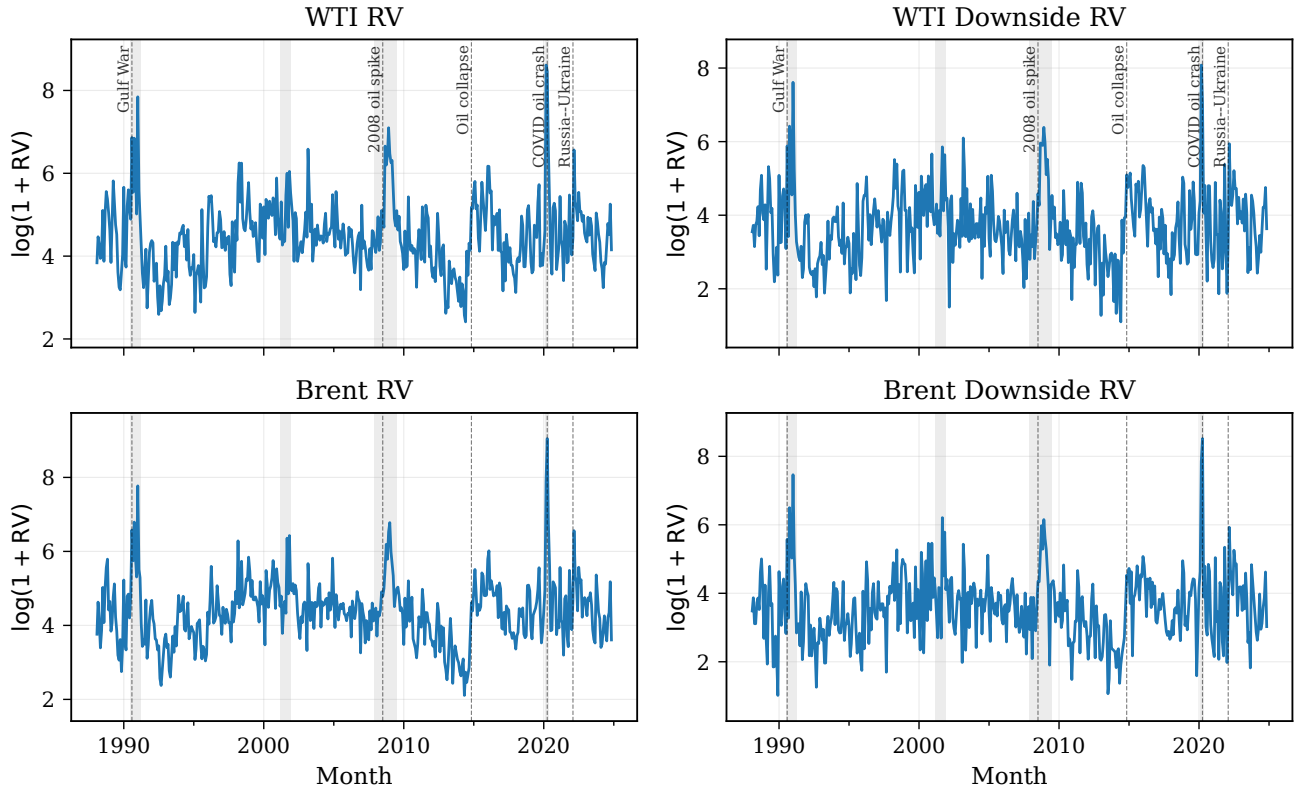


Figure 2: Monthly Oil-Price Realized Volatility

Notes: Monthly realized volatility and downside realized volatility for WTI and Brent, measured as $\log(1 + RV)$ and $\log(1 + DRV)$. Gray bands represent NBER recessions; dashed lines mark selected oil-market events.

Table 2
Summary statistics

Variable	Mean	SD	P10	Median	P90	Min	Max
<i>Panel A: Oil narrative indices</i>							
Oil Attention	0.003	0.995	-1.285	-0.009	1.256	-3.144	3.653
Oil Risk Attention	0.004	0.995	-1.299	-0.040	1.347	-2.602	3.755
Oil Uncertainty	0.006	0.998	-1.130	-0.096	1.349	-2.757	3.818
Oil Tone	0.000	0.998	-0.996	-0.226	1.244	-1.668	4.761
<i>Panel B: Oil-market risk outcomes</i>							
WTI RV	4.466	0.866	3.485	4.409	5.468	2.413	8.613
WTI Downside RV	3.709	0.984	2.534	3.672	4.896	1.101	8.080
Brent RV	4.365	0.858	3.362	4.361	5.308	2.108	9.048
Brent Downside RV	3.574	0.994	2.331	3.573	4.772	1.018	8.517

Notes: Summary statistics for the monthly variables in the baseline regressions. Panel A indices are standardized over the estimation sample. Panel B outcomes are $\log(1 + RV)$ and $\log(1 + DRV)$. The sample is $T = 442$ monthly forecast origins, February 1988 through November 2024.

3. Empirical Design

This section sets out the empirical design: predictive regressions and a recursive out-of-sample framework for forecasting one-month-ahead oil-price volatility from oil-news coverage.

3.1. Predictive Regressions

The baseline predictive regression is

$$y_{t+1} = \alpha + \rho y_t + \beta \text{Narrative}_t + \varepsilon_{t+1}, \quad (8)$$

where y_{t+1} is $\log(1 + RV_{t+1})$ or $\log(1 + DRV_{t+1})$ and y_t controls for volatility persistence, and Narrative_t is one of the standardized text measures. Because the measures are standardized, β is the response to a one-standard-deviation change in the predictor, and a positive β means that heavier oil coverage in month t predicts higher risk in month $t + 1$. Realized volatility is persistent and serially correlated, so inference uses Newey and West (1987) standard errors with six lags. The regression is estimated separately for each of

the four outcomes,

$$y_{t+1}^i \in \{\log(1 + RV_{t+1}^i), \log(1 + DRV_{t+1}^i)\}, \quad (9)$$

$$i \in \{WTI, Brent\}.$$

3.2. Out-of-Sample Forecasting

In-sample association is a first step, but the main task is out-of-sample forecasting. We begin with an autoregressive realized-volatility model,

$$y_{t+1} = \alpha + \rho y_t + u_{t+1}, \quad (10)$$

augmented with a text measure,

$$y_{t+1} = \alpha + \rho y_t + \beta \text{Narrative}_t + u_{t+1}. \quad (11)$$

Forecasts are recursive on an expanding window: at each origin, the model uses only data available up to that point. The out-of-sample period begins in January 2000 and ends with the November 2024 origin, which forecasts December 2024, giving 299 forecasts.

We then consider the HAR model, which captures persistence at several horizons (Corisi, 2009),

$$y_{t+1} = \alpha + \beta_1 y_t + \beta_2 \bar{y}_t^{(3)} + \beta_3 \bar{y}_t^{(12)} + u_{t+1}, \quad (12)$$

augmented with the text measure,

$$y_{t+1} = \alpha + \beta_1 y_t + \beta_2 \bar{y}_t^{(3)} + \beta_3 \bar{y}_t^{(12)} + \gamma \text{Narrative}_t + u_{t+1}, \quad (13)$$

where $\bar{y}_t^{(3)}$ and $\bar{y}_t^{(12)}$ are the trailing three- and twelve-month averages available at the origin. Beating HAR is harder than beating AR, which makes it the more demanding benchmark.

3.3. Forecast Evaluation

Forecasts are compared by out-of-sample mean squared error. With $\hat{y}_{t+1}^B$ the benchmark forecast and $\hat{y}_{t+1}^M$ the augmented forecast, the out-of-sample R^2 is

$$R_{OOS}^2 = 1 - \frac{\sum_{t \in \mathcal{O}} (y_{t+1} - \hat{y}_{t+1}^M)^2}{\sum_{t \in \mathcal{O}} (y_{t+1} - \hat{y}_{t+1}^B)^2}, \quad (14)$$

which is positive when the augmented model has the smaller error. Because the augmented models nest the benchmarks, inference uses the Clark and West (2007) MSPE-adjusted test, which corrects for the upward bias the larger model incurs from estimating an extra parameter under the null. The adjusted loss differential is

$$f_{t+1} = (y_{t+1} - \hat{y}_{t+1}^B)^2 - [(y_{t+1} - \hat{y}_{t+1}^M)^2 - (\hat{y}_{t+1}^B - \hat{y}_{t+1}^M)^2], \quad (15)$$

whose positive mean indicates improvement, tested with Newey and West (1987) standard errors. Applied to a model that already contains one text measure, the same statistic tests whether a second measure adds anything; this is the encompassing test behind Section 4.3. We also plot cumulative gains,

$$\text{CumGain}_\tau = \sum_{t \leq \tau} [(y_{t+1} - \hat{y}_{t+1}^B)^2 - (y_{t+1} - \hat{y}_{t+1}^M)^2], \quad (16)$$

to show whether the gains are broad or episodic. Finally, because the paper's claim concerns risk rather than fit, Section 4.4 evaluates the forecasts on the lower tail of the return distribution itself, scoring Value-at-Risk with the quantile loss and the joint VaR-and-ES forecasts with the consistent loss of Fissler and Ziegel (2016).

4. Main Results

This section presents the main results. Oil-news attention is informative about future oil-price volatility, with the largest gains for downside risk. Two further conclusions follow. First, the forecasting power comes from the volume of news coverage rather than its risk content: risk-weighted coverage outperforms plain volume against volatility-history benchmarks, but the difference disappears once option-implied volatility is controlled for. Second, the predictive content extends to the lower quantiles of the return distribution, the quantity that matters most for risk management.

4.1. Predictive Regressions

Table 3 reports regressions of next-month realized volatility on current realized volatility and each narrative measure in turn, for WTI and Brent and for both total and downside realized volatility. Realized volatility enters in logs and each measure is standardized, so each coefficient gives the approximate proportional change in next-month volatility associated with a one-standard-deviation increase in the predictor.

Oil-news coverage predicts next-month volatility. Conditional on current volatility, a one-standard-deviation increase in coverage in month t predicts realized volatility in month $t+1$ that is roughly 9 to 10 percent higher for WTI and Brent, and the effect on downside realized volatility, roughly 18 to 20 percent, is nearly twice as large. Risk attention yields coefficients of similar size and significance because, as Section 4.3 shows, the two measures are nearly the same series.

The other two measures are weaker. Oil uncertainty predicts downside but not total realized volatility, consistent with uncertainty language concentrating in adverse news. Oil tone predicts neither outcome for either crude. This last result is worth emphasizing: the directional content of oil-news language, which bears on where the price is headed, carries no predictive information about how much it will move.

4.2. Out-of-Sample Forecasts

The predictive regressions establish an in-sample relation. The more demanding question is whether attention improves forecasts out of sample. Tables 4 and 5 report the out-of-sample R^2 relative to two benchmarks, an autoregressive realized-volatility model and the HAR model, which captures volatility persistence at monthly, quarterly, and annual horizons. Forecasts are recursive on an expanding estimation window, with the first forecast in January 2000.

Relative to the AR benchmark (Table 4), oil-news attention improves forecasts of all four outcomes, with the

Table 3
Predictive regressions of realized volatility on oil narrative indices

Narrative measure	WTI		Brent	
	RV	Downside RV	RV	Downside RV
Oil Attention	0.0932*** (2.72)	0.1762*** (3.57)	0.1034*** (3.08)	0.1974*** (4.15)
Δ Adj. R^2	0.008	0.026	0.011	0.032
Oil Risk Attention	0.0961*** (2.89)	0.1925*** (4.03)	0.0976*** (2.98)	0.2133*** (4.71)
Δ Adj. R^2	0.008	0.030	0.009	0.037
Oil Uncertainty	0.0244 (1.01)	0.0676** (2.07)	0.0220 (0.91)	0.0925*** (2.67)
Δ Adj. R^2	-0.001	0.003	-0.001	0.007
Oil Tone	-0.0291 (-0.79)	0.0045 (0.09)	-0.0262 (-0.77)	0.0204 (0.43)
Δ Adj. R^2	0.000	-0.002	0.000	-0.001
Baseline AR Adj. R^2	0.397	0.200	0.430	0.169
Observations	442	442	442	442

Notes: Each column block reports a separate predictive regression of one-month-ahead realized volatility (RV) or downside realized volatility on current RV and one standardized oil-narrative index. Newey–West t -statistics with six lags (Newey and West, 1987) are in parentheses. Δ Adj. R^2 is the change in adjusted R^2 relative to the autoregressive benchmark, whose adjusted R^2 is reported in the bottom panel. The sample spans February 1988 to November 2024 ($T = 442$). $*$ $p < 0.10$, $**$ $p < 0.05$, $***$ $p < 0.01$.

largest gains for downside realized volatility: an R^2_{OOS} of 4.2 percent for WTI and 4.7 percent for Brent. Weighting coverage by risk language raises the gains in three of the four columns, a distinction we study in Section 4.3. Against the more demanding HAR benchmark (Table 5) the gains persist and again concentrate in the downside, at 3.7 percent for WTI and 4.7 percent for Brent, with risk attention larger than plain attention in every column. The substantive point is that attention adds forecasting power on top of HAR: coverage volume is not merely another measure of past volatility but carries information that lagged volatility at three horizons does not.

These gains are not the product of one or two episodes. Figures 3 and 4 plot the cumulative gains over the out-of-sample period. For oil-news coverage they are positive and rising throughout the sample, with the clearest improvements during the 2014–2016 price collapse, the 2020 crash, and the post-2020 period of elevated geopolitical risk.

A related question is whether particular types of oil news carry the predictive signal. Classifying oil-centered text into themes (geopolitics, OPEC, inventories, demand, macro, refining, and energy transition), we find that demand, macro, geopolitical, and inventory news account for much of the signal, with the same downside tilt. We report these results in Appendix C.4 and read them as descriptive, because the themes share vocabulary and are not distinct channels.

4.3. Is the Signal Attention or Risk Content?

Attention and risk attention produce almost identical coefficients in Table 3 and very similar out-of-sample gains in Tables 4 and 5, because the two indices are nearly the same series: they correlate at 0.93. In a paper that builds a risk

dictionary, it is natural to ask whether weighting coverage by risk language adds anything over plain coverage. Table 6 addresses this in-sample and out-of-sample.

Panel A puts both indices in the same regression. For total realized volatility, neither survives, because the two indices are too highly correlated for the regression to tell them apart. For downside volatility the picture is sharper: the coefficient on attention collapses toward zero while risk attention holds its size and remains marginally significant.

The out-of-sample evidence in Panel B points the same way. In nested encompassing tests, adding risk attention to a HAR model that already contains attention improves downside-volatility forecasts by 1.1 percentage points for WTI and 1.3 for Brent, both significant at the 5% level, whereas adding attention to a model that already contains risk attention never improves forecasts. For total volatility, neither measure adds to the other. Figure 5 shows the asymmetry building over time: the gain from adding risk attention rises steadily, while the gain from adding attention stays flat.

So far this looks like a win for the risk dictionary, but the advantage is fragile. Risk attention outperforms plain attention only against volatility-history benchmarks, and the edge disappears once option-implied volatility enters the regression. Table 7 makes the point directly: against the HAR benchmark, risk attention adds 1.1 and 1.3 percentage points to the downside forecast, but once the VIX or the OVX is included, the increment falls to 0.2 points or less and is insignificant, with the largest t -statistic equal to 1.0 (full detail in Appendix B.4). Whatever distinguishes risk-focused coverage from general coverage is already captured by option-implied volatility, and conditional on it the two indices convey the same forecasting information. Because

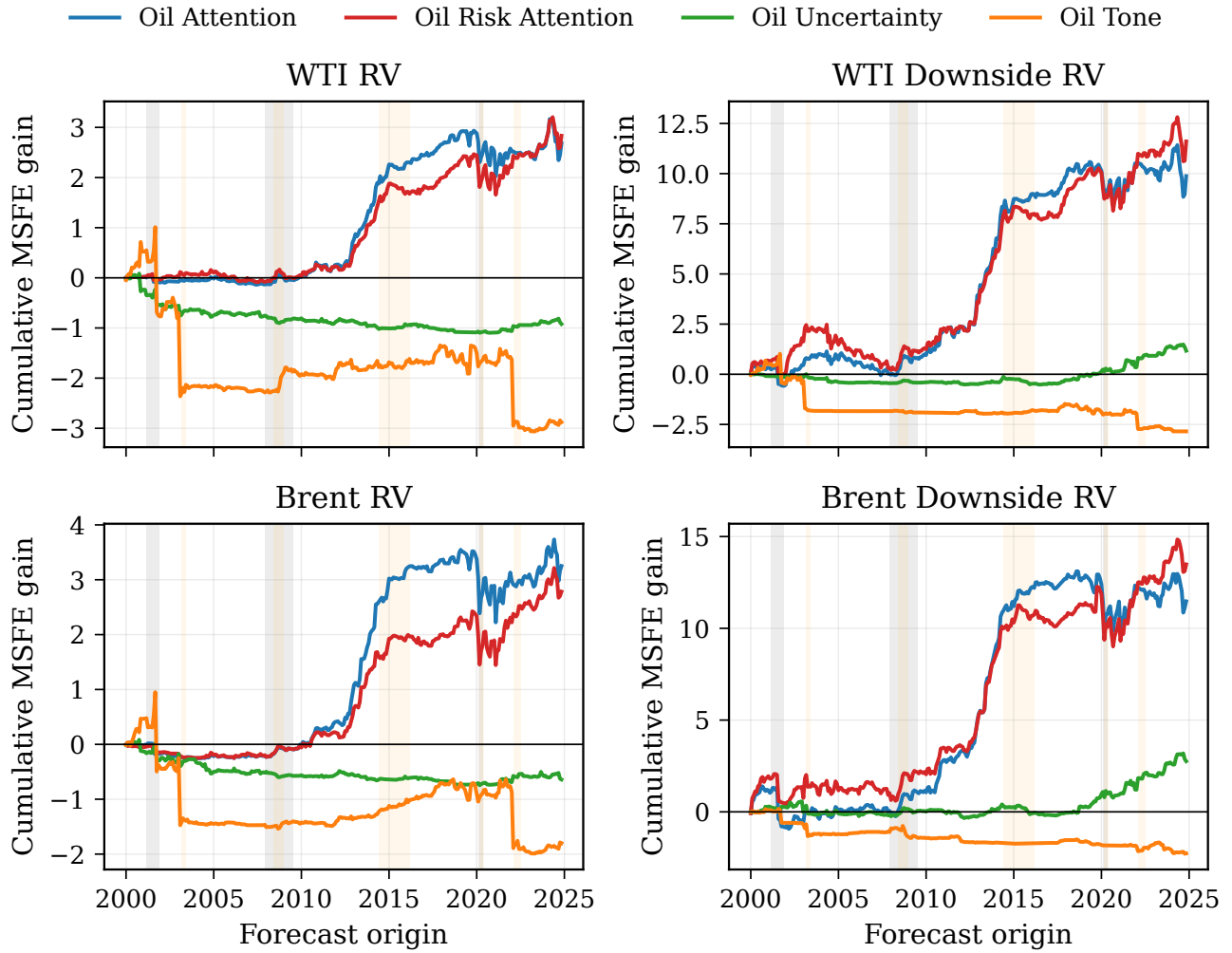


Figure 3: Cumulative Out-of-Sample Forecast Gains, AR Benchmark

Notes: Cumulative MSFE gains of each narrative-augmented model relative to the AR benchmark. A positive value means the augmented model has produced lower cumulative squared forecast errors than the benchmark. Gray bands denote NBER recessions; light orange bands mark selected oil-market episodes.

Table 4
Out-of-sample forecasting performance against the AR benchmark (% R^2_{OOS})

Narrative measure	WTI		Brent	
	RV	Downside RV	RV	Downside RV
Oil Attention	2.07*** (3.07)	4.15*** (3.64)	2.61*** (2.98)	4.67*** (3.47)
Oil Risk Attention	2.18*** (3.30)	4.87*** (4.10)	2.23*** (2.97)	5.48*** (4.09)
Oil Uncertainty	-0.71 (-1.30)	0.49** (1.65)	-0.51 (-1.15)	1.12** (2.30)
Oil Tone	-2.22 (-0.35)	-1.20 (-1.09)	-1.44 (-0.46)	-0.92 (-1.65)
OOS observations	299	299	299	299

Notes: Out-of-sample R^2_{OOS} in percentage points relative to an AR benchmark, following Campbell and Thompson (2008). Positive values indicate a lower mean squared forecast error than the benchmark. Clark–West MSPE-adjusted statistics (Clark and West, 2007) are in parentheses; one-sided p -values test the null that the benchmark is at least as accurate. Forecasts are one step ahead from an expanding window; the evaluation sample runs from January 2000 to November 2024 ($T = 299$). $\dot{p} < 0.10$, $\dot{p}^* < 0.05$, $\dot{p}^{**} < 0.01$.

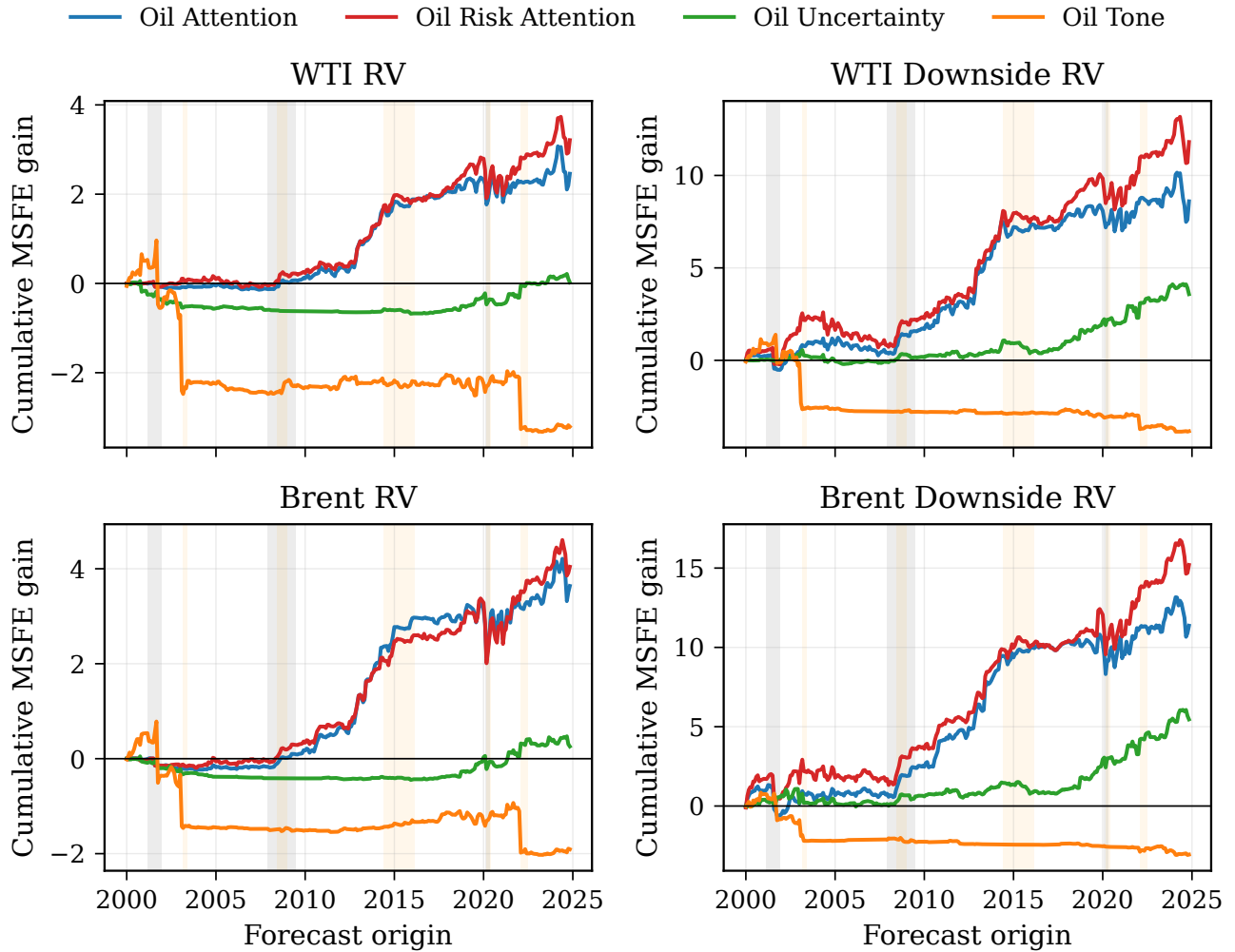


Figure 4: Cumulative Out-of-Sample Forecast Gains, HAR Benchmark

Notes: Cumulative MSFE gains of each narrative-augmented model relative to the HAR benchmark. A positive value means the augmented model has produced lower cumulative squared forecast errors than the benchmark. Gray bands denote NBER recessions; light orange bands mark selected oil-market episodes.

plain coverage volume is the simpler measure, it is the natural object of interest: the risk dictionary remains useful descriptively but adds no forecasting information once market-based controls are included. From here on we treat oil-news attention as the object of interest and report risk attention alongside it as a near-equivalent.

4.4. Forecasting the Tail: Value-at-Risk

The forecasting tables evaluate volatility forecasts by squared error, but what matters to a risk manager is the lower tail of the return distribution. This subsection therefore evaluates the forecasts on the tail directly: each volatility forecast is translated into a one-month-ahead Value-at-Risk (VaR) and Expected Shortfall (ES) forecast and scored with loss functions designed for the tail.

At each forecast origin, the volatility forecast supplies the conditional standard deviation of next-month returns, from which we compute the 5% and 1% VaR and ES. Because a Gaussian conditional distribution understates the

extreme tail of oil returns, we consider three distributions: a Gaussian benchmark; a Student- t with five degrees of freedom, scaled to unit variance; and filtered historical simulation (FHS), which takes the tail from the empirical distribution of the model's standardized forecast residuals (Kuester, Mittnik and Paolella, 2006). We compare two benchmark volatility models, HAR and HAR augmented with the oil-specific OVX (forecasts from 2010 onward), against the same models with oil-news attention added. VaR is scored with the quantile (tick) loss, a consistent scoring function for a quantile, and the joint VaR–ES forecast with the Fissler and Ziegel (2016) loss in the form proposed by Patton, Ziegel and Chen (2019). Models are compared with Diebold and Mariano (1995) tests on the loss differentials.

Table 8 reports the calibration of the HAR benchmark under each distribution. The Gaussian gives a well-calibrated 5% VaR for WTI and a marginal one for Brent, but it understates the 1% tail: violation frequencies run

Table 5Out-of-sample forecasting performance against the HAR benchmark (% R^2_{OOS})

Narrative measure	WTI		Brent	
	RV	Downside RV	RV	Downside RV
Oil Attention	1.88*** (3.11)	3.70*** (3.80)	2.91*** (3.45)	4.71*** (4.21)
Oil Risk Attention	2.46*** (3.51)	5.08*** (4.51)	3.23*** (3.61)	6.30*** (4.98)
Oil Uncertainty	0.02 (0.48)	1.54*** (2.68)	0.21 (0.83)	2.26*** (3.14)
Oil Tone	-2.46 (-0.48)	-1.65 (-0.81)	-1.52 (-0.63)	-1.27 (-1.33)
OOS observations	299	299	299	299

Notes: Out-of-sample R^2_{OOS} in percentage points relative to a HAR benchmark that includes current realized volatility and its trailing three- and twelve-month averages. Positive values indicate a lower mean squared forecast error than the benchmark. Clark–West MSPE-adjusted statistics (Clark and West, 2007) are in parentheses; one-sided p -values test the null that the benchmark is at least as accurate. Forecasts are one step ahead from an expanding window; the evaluation sample runs from January 2000 to November 2024 ($T = 299$). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 6

Attention versus risk attention: incremental predictive content

	WTI		Brent	
	RV	Downside RV	RV	Downside RV
<i>Panel A: In-sample predictive regressions</i>				
Oil Attention (alone)	0.093***	0.176***	0.103***	0.197***
Oil Risk Attention (alone)	0.096***	0.193***	0.098***	0.213***
Joint specification				
Oil Attention	0.042 (0.58)	0.010 (0.10)	0.091 (1.20)	0.023 (0.21)
Oil Risk Attention	0.058 (0.84)	0.183* (1.91)	0.014 (0.19)	0.192* (1.80)
<i>Panel B: Out-of-sample encompassing, HAR benchmark (R^2_{OOS}, pp)</i>				
Risk added to HAR + Attention	0.2 (1.09)	1.1** (2.26)	0.0 (0.37)	1.3** (2.44)
Attention added to HAR + Risk	-0.4 (-1.40)	-0.3 (-0.26)	-0.4 (-0.75)	-0.3 (-0.38)

Notes: Panel A reports predictive regressions of next-month $\log(1 + \text{RV})$ or $\log(1 + \text{DRV})$ on current realized volatility and the indicated standardized measure(s). The “alone” rows repeat the estimates in Table 3, where t -statistics are reported; the joint specification includes both measures. Newey–West t -statistics with six lags (Newey and West, 1987) are in parentheses. Oil attention and oil risk attention have a correlation of 0.93. Panel B reports the out-of-sample R^2 (percentage points) from adding one measure to a HAR model that already includes the other, with Clark–West MSPE-adjusted statistics (Clark and West, 2007) in parentheses. Forecasts are one step ahead from an expanding window over $T = 299$ origins (2000:01–2024:11). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

about three to four times the nominal level, and the Kupiec (1995) test rejects the 1% VaR for both crudes. The heavier-tailed specifications help. Filtered historical simulation is well calibrated at both the 5% and the 1% level for both crudes, and the 1% WTI VaR under the Student- t is no longer rejected. At the 5% level, the Student- t is somewhat less well calibrated than the Gaussian, as expected, since its fatter tails pull the 5% quantile inward; the 1% Brent VaR remains miscalibrated under the Gaussian and the Student- t . Because this calibration is of the benchmark model, which uses no

text information, it reflects a limitation of monthly volatility forecasting rather than of the attention measure.

Table 9 shows whether adding oil-news attention improves tail-forecast accuracy. At the 5% level, attention improves the VaR quantile score against the OVX-augmented benchmark under all three distributions, with Diebold and Mariano (1995) statistics between 2.8 and 4.0; it also improves the VaR against the plain HAR benchmark in five of the six cases. For Expected Shortfall, the evidence points the same way against the OVX-augmented benchmark, where

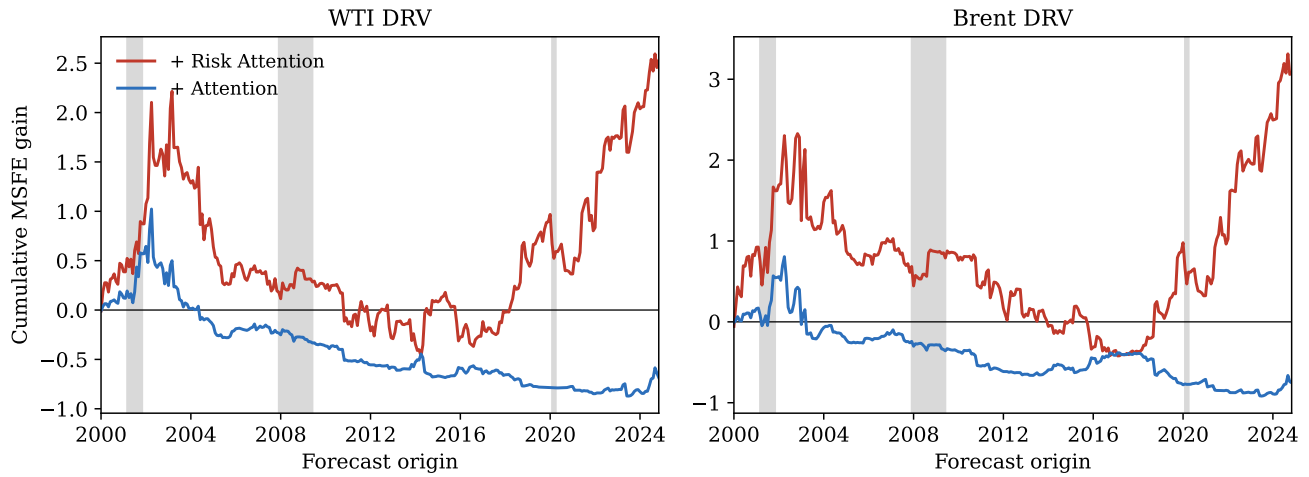


Figure 5: Incremental Forecast Gain of Each Measure on Downside Realized Volatility

Notes: Cumulative out-of-sample MSFE gain on downside realized volatility from adding each narrative measure to a HAR model that already contains the other. The red line adds risk attention to HAR+attention; the blue line adds attention to HAR+risk attention. A rising line indicates that the added measure improves on the model without it. Gray bands denote NBER recessions.

Table 7

Is the incremental signal attention or risk content? Robustness to implied volatility

Benchmark (already contains attention)	WTI	Brent
HAR	1.1** (2.26)	1.3** (2.44)
HAR + VIX	0.2 (1.00)	0.2 (0.87)
HAR + OVX	-0.3 (0.24)	-0.2 (0.15)

Notes: Out-of-sample R^2_{OOS} (percentage points) on downside realized volatility from adding risk attention to a benchmark that already contains plain oil-news attention and, where indicated, an implied-volatility index. A positive, significant value would mean that weighting coverage by risk language forecasts beyond coverage volume and option-implied volatility. Clark and West (2007) MSPE-adjusted statistics in parentheses. Against the volatility-history benchmark (HAR) risk attention adds 1.1 (WTI) and 1.3 (Brent) percentage points, both significant; once the VIX or the OVX is included the increment is statistically zero. The full set of benchmarks, including the AR specifications and total realized volatility, is in Appendix B.4 (Table B.6). $\hat{p} < 0.10$, $\hat{p}^* < 0.05$, $\hat{p}^{**} < 0.01$.

the improvement is significant for both crudes under all three distributions; against the plain HAR benchmark, the ES improvement is significant only under filtered historical simulation. We do not report the 1% comparison, where power is very low, with only about three violations expected over the sample. The VaR improvements are thus not an artifact of the normality assumption: they hold under fat-tailed and fully nonparametric distributions, and they hold even against a benchmark that already includes oil-specific implied volatility.

Two caveats keep the exercise honest. First, the VaR and ES are obtained by scaling a volatility forecast by a fixed distribution rather than by modeling the conditional

quantile directly. Second, as Table 8 shows, the 1% Brent tail is miscalibrated under the Gaussian and the Student- t (though not under filtered historical simulation). Neither caveat affects the comparison in Table 9, which holds the distribution fixed and varies only whether attention is added to the volatility forecast. The gains from oil-news attention thus extend beyond squared-error loss: they hold for the 5% VaR under parametric and nonparametric tails, and for ES against the OVX-augmented benchmark.

4.5. Why Coverage Volume Forecasts Risk

Interpreting the findings is straightforward. Periods of heavy oil coverage are periods of fragility in the oil market, and that fragility shows up as larger price movements, particularly on the downside. Attention measures the spread of the return distribution rather than its direction, and it forecasts both that spread and its tail. It does so out of sample, beyond the volatility dynamics captured by the HAR model and beyond option-implied volatility, and under loss functions that focus directly on the tail.

Two further results pinpoint the source of this predictive content. First, the directional content of news, measured by the tone index, does not help forecast volatility, either in-sample or out-of-sample. Second, weighting coverage by risk language helps relative to plain volume and to benchmarks built from past volatility (Section 4.3), but that advantage disappears once option-implied volatility is included. What remains is the volume of coverage, which raises the question of why a count of articles should matter for risk.

The mixture-of-distributions hypothesis provides an answer (Clark, 1973; Tauchen and Pitts, 1983; Andersen, 1996). Under this hypothesis, returns are drawn from a mixture of distributions whose scale rises with the rate at which new information arrives, so periods of more frequent information arrival are periods of higher volatility. In our setting, oil coverage is an observable proxy for that arrival

Table 8

Calibration of the benchmark Value-at-Risk under alternative distributions

Asset	α	Gaussian		Student- $t(5)$		FHS	
		Hit rate (%)	Kupiec p	Hit rate (%)	Kupiec p	Hit rate (%)	Kupiec p
WTI	5%	5.0	0.98	6.7	0.20	5.0	0.98
	1%	3.0	<0.01	2.0	0.12	1.0	0.99
Brent	5%	7.4	0.08	8.1	0.03	5.0	0.98
	1%	4.0	<0.01	3.0	<0.01	1.0	0.99

Notes: Realized violation frequency (hit rate) of the one-month-ahead HAR Value-at-Risk at nominal level α , and the p -value of the Kupiec (1995) unconditional-coverage test. The conditional distributions are Gaussian, a unit-variance Student- t with five degrees of freedom, and filtered historical simulation (FHS). Forecasts use expanding-window estimation over the out-of-sample period, 2000:01–2024:11 ($T = 298$). A p -value below 0.05 rejects correct coverage.

Table 9

Value-at-Risk and Expected Shortfall: improvement from adding oil-news attention

Asset	Benchmark	Gaussian	Student- $t(5)$	FHS
<i>Panel A: Value-at-Risk (quantile-loss DM statistic, $\alpha = 5\%$)</i>				
WTI	HAR	2.63***	1.83*	0.92
	HAR + OVX	2.83***	2.82***	3.98***
Brent	HAR	2.41**	2.18**	2.58***
	HAR + OVX	3.12***	3.53***	3.21***
<i>Panel B: Expected Shortfall (FZ-loss DM statistic, $\alpha = 5\%$)</i>				
WTI	HAR	1.19	0.97	1.69*
	HAR + OVX	2.90***	3.03***	2.79***
Brent	HAR	0.52	0.45	1.72*
	HAR + OVX	2.34**	2.50**	2.77***

Notes: Diebold and Mariano (1995) (DM) t -statistics for the difference in loss between the indicated benchmark volatility model and the same model augmented with oil-news attention, at the 5% tail. Panel A uses the quantile (tick) loss for Value-at-Risk; Panel B uses the joint Value-at-Risk and Expected Shortfall loss of Fissler and Ziegel (2016) (FZ). Positive values indicate that adding attention lowers the loss. Columns vary the conditional distribution: Gaussian, unit-variance Student- $t(5)$, and filtered historical simulation (FHS). The HAR rows use the 2000:01–2024:11 out-of-sample period (298 backtest observations); the HAR+OVX rows use the shorter sample for which the OVX is available (OVX begins June 2007), with the OVX-conditioned expanding-window forecasts beginning in 2010 (174 backtest observations). Two-sided tests: $\hat{p} < 0.10$, $\hat{p}^* < 0.05$, $\hat{p}^{**} < 0.01$.

rate: OPEC decisions, geopolitical developments, supply disruptions, and demand shifts each generate news, and months with more such events produce both more coverage and larger price movements.

This interpretation is consistent with all three patterns. First, it explains why the volume of coverage, rather than its tone, carries the signal: tone conveys the sign of the news, whereas volatility responds to the arrival of information regardless of sign. Second, it explains why weighting coverage by risk language adds little: the risk-weighted measure is largely a rescaling of the same arrival count, and whatever it adds is already embedded in option-implied volatility. Third, together with the assumption that adverse oil-market news attracts disproportionate coverage, it explains why the lower tail is the most informative part of the distribution. Coverage volume thus forecasts the skewness of the return distribution, a thicker left tail, without forecasting the sign of the expected return. Read this way, the fact that the predictor is ultimately a count of articles is not a weakness of the measure: counting

information arrivals is precisely what the mechanism says should matter.

A final question is why option prices do not already incorporate this observable count of news. Implied volatility is a risk-neutral expectation that embeds a time-varying variance risk premium (Bollerslev, Tauchen and Zhou, 2009), so it is a noisy forecast of realized volatility even when option markets are informationally efficient. Coverage volume measures the arrival rate directly and embeds no risk premium, so it can carry forecasting information that implied volatility obscures. We do not test this channel directly.

5. Robustness and Additional Evidence

This section summarizes a set of robustness checks and additional evidence on the economic relevance of our results. All supporting results are reported in the Online Appendix.

First, we assess whether the findings are sensitive to how the newspaper measures are constructed (Appendix A). The forecasting gains are robust to the transformation applied to the monthly article count (Appendix A.2), to the

width of the contextual window around oil-related text (Appendix A.3), and to the composition of the newspaper sample (Appendix A.4); the signal is present in all four newspapers individually. Risk-and-uncertainty language retains its predictive content for downside volatility whether it is normalized by the number of oil-related articles or by the number of words in the oil-centered windows (Appendix A.5).

Second, we examine the stability of the forecasting gains. They continue to hold across alternative forecast-evaluation start dates, after excluding the March–June 2020 oil-market disruption, and under a more stringent corpus that removes recurring market digests and applies a stricter oil-relevance threshold (Appendix B.1). They are also robust across subsamples spanning the transition from print to online news (Appendix B.2).

Third, we ask whether news attention is merely a proxy for observable market information or for well-known risk indicators. Adding the monthly oil return and its absolute value to the regressions leaves the attention measures significant for all four volatility outcomes (Appendix B.3). Controlling for the VIX and the OVX in the benchmark specifications, aggregate oil-news attention remains economically and statistically significant for downside realized volatility (Appendix B.4). These encompassing tests also clarify the decomposition in Section 4.3: risk attention loses predictive power once implied volatility is controlled for, whereas aggregate oil-news attention retains it. News attention also keeps its information content when we control jointly for the Economic Policy Uncertainty and Geopolitical Risk indices (Appendix B.5).

Finally, Appendix C reports several extensions. In a forecast horse race, aggregate attention and risk attention deliver the largest gains for downside realized volatility, and adding attention to a combination of conventional predictors improves the combined forecast (Appendix C.1). The forecasts also generate economic value in a volatility-timing exercise (Appendix C.2). Attention predicts both upside and downside realized volatility, consistent with oil-market news generating large price movements in either direction, though the gains are larger and more precisely estimated for the downside (Appendix C.3).

6. Conclusion

This paper asks whether the volume of oil-market news coverage in leading U.S. newspapers helps forecast oil-price volatility out of sample. Adding this oil-news attention to autoregressive and HAR benchmarks improves forecasts of realized volatility for WTI and Brent, with the largest gains for downside realized volatility. The improvement survives benchmarks that already include option-implied volatility, from either the VIX or the OVX, and it extends to the tail: attention improves 5% Value-at-Risk forecasts under Gaussian, Student-*t*, and filtered historical simulation distributions, even against a benchmark conditioned on the OVX.

The predictive content of oil-news attention comes from the volume of coverage, not from its sentiment or risk

content. The tone index does not forecast volatility, and the advantage of risk-weighted coverage over volatility-history benchmarks does not survive the inclusion of implied volatility. For forecasting risk, what matters in oil coverage is the salience of the news, not its sentiment. This pattern is consistent with the mixture-of-distributions hypothesis, under which the volume of coverage proxies for the rate at which information about the oil market arrives, and volatility rises with that arrival rate regardless of the direction of the news.

Two extensions suggest themselves. First, we can evaluate only the 5% Value-at-Risk, because the 1% tail yields too few violations in monthly data; resolving the deeper tail would require daily or weekly coverage. Second, the analysis could be extended to oil futures markets and to spillovers from oil-market volatility to energy-equity markets. More broadly, the results show that simply counting oil-market news stories forecasts oil-price volatility, and its downside tail in particular.

Data availability

The raw newspaper text was obtained from ProQuest TDM Studio and cannot be redistributed under ProQuest's terms of use. All derived data (the monthly narrative indices) and the replication code are available at <https://github.com/gabeabeyie/oil-tail-risk-newspaper-attention>.

Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work, the author used Anthropic's Claude in order to improve the readability and language of the manuscript. After using this tool, the author reviewed and edited the content as needed and takes full responsibility for the content of the publication.

Acknowledgements

I thank the Editor, Julien Chevallier, and three anonymous referees for comments that substantially improved the paper. I am grateful to Andrew Hanson, Luiz Lima, and Mohammed Mohsin, and to participants in the brownbag seminar at the University of Tennessee Department of Economics, for helpful discussions. Any remaining errors are my own.

Declaration of competing interest

The author declares that he has no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- Abeyie, G.A., 2026. Beyond news headlines and tf-idf: Enhancing text-based forecasting models with validated collocations and improved attention. *International Journal of Forecasting*.

- Alquist, R., Kilian, L., Vigfusson, R.J., 2013. Forecasting the price of oil, in: Elliott, G., Timmermann, A. (Eds.), *Handbook of Economic Forecasting*. Elsevier, volume 2, pp. 427–507.
- Andersen, T.G., 1996. Return volatility and trading volume: An information flow interpretation of stochastic volatility. *The Journal of Finance* 51, 169–204.
- Andersen, T.G., Bollerslev, T., Diebold, F.X., Labys, P., 2003. Modeling and forecasting realized volatility. *Econometrica* 71, 579–625.
- Andrei, D., Hasler, M., 2015. Investor attention and stock market volatility. *Review of Financial Studies* 28, 33–72.
- Baker, S.R., Bloom, N., Davis, S.J., 2016. Measuring economic policy uncertainty. *Quarterly Journal of Economics* 131, 1593–1636.
- Barbaglia, L., Consoli, S., Manzan, S., 2023. Forecasting with economic news. *Journal of Business & Economic Statistics* 41, 708–719.
- Baumeister, C., Hamilton, J.D., 2019. Structural interpretation of vector autoregressions with incomplete identification: Revisiting the role of oil supply and demand shocks. *American Economic Review* 109, 1873–1910.
- Baumeister, C., Kilian, L., 2015. Forecasting the real price of oil in a changing world: A forecast combination approach. *Journal of Business & Economic Statistics* 33, 338–351.
- Bollerslev, T., Li, S.Z., Zhao, B., 2020. Good volatility, bad volatility, and the cross section of stock returns. *Journal of Financial and Quantitative Analysis* 55, 751–781.
- Bollerslev, T., Tauchen, G., Zhou, H., 2009. Expected stock returns and variance risk premia. *The Review of Financial Studies* 22, 4463–4492.
- Bybee, L., Kelly, B.T., Manela, A., Xiu, D., 2024. Business news and business cycles. *Journal of Finance* 79, 3105–3147.
- Caldara, D., Iacoviello, M., 2022. Measuring geopolitical risk. *American Economic Review* 112, 1194–1225.
- Campbell, J.Y., Thompson, S.B., 2008. Predicting excess stock returns out of sample: Can anything beat the historical average? *Review of Financial Studies* 21, 1509–1531.
- Clark, P.K., 1973. A subordinated stochastic process model with finite variance for speculative prices. *Econometrica* 41, 135–155.
- Clark, T.E., West, K.D., 2007. Approximately normal tests for equal predictive accuracy in nested models. *Journal of Econometrics* 138, 291–311.
- Corsi, F., 2009. A simple approximate long-memory model of realized volatility. *Journal of Financial Econometrics* 7, 174–196.
- Da, Z., Engelberg, J., Gao, P., 2011. In search of attention. *Journal of Finance* 66, 1461–1499.
- Diebold, F.X., Mariano, R.S., 1995. Comparing predictive accuracy. *Journal of Business and Economic Statistics* 13, 253–263.
- Fissler, T., Ziegel, J.F., 2016. Higher order elicibility and osband's principle. *The Annals of Statistics* 44, 1680–1707.
- Garcia, D., 2013. Sentiment during recessions. *Journal of Finance* 68, 1267–1300.
- Hamilton, J.D., 2009. Causes and consequences of the oil shock of 2007–08. *Brookings Papers on Economic Activity* Spring, 215–261.
- Hong, Y., Jiang, F., Meng, L., Xue, B., 2025. Forecasting inflation using economic narratives. *Journal of Business & Economic Statistics* 43, 216–231.
- Jurado, K., Ludvigson, S.C., Ng, S., 2015. Measuring uncertainty. *American Economic Review* 105, 1177–1216.
- Kilian, L., 2009. Not all oil price shocks are alike: Disentangling demand and supply shocks in the crude oil market. *American Economic Review* 99, 1053–1069.
- Kuester, K., Mittnik, S., Paolella, M.S., 2006. Value-at-risk prediction: A comparison of alternative strategies. *Journal of Financial Econometrics* 4, 53–89.
- Kupiec, P.H., 1995. Techniques for verifying the accuracy of risk measurement models. *The Journal of Derivatives* 3, 73–84.
- Li, X., Shang, W., Wang, S., 2019. Text-based crude oil price forecasting: A deep learning approach. *International Journal of Forecasting* 35, 1548–1560.
- Li, Y., Jiang, S., Li, X., Wang, S., 2021. The role of news sentiment in oil futures returns and volatility forecasting: Data-decomposition based deep learning approach. *Energy Economics* 95, 105140.
- Lima, L.R., Godeiro, L.L., 2023. Equity-premium prediction: Attention is all you need. *Journal of Applied Econometrics* 38, 105–122.
- Loughran, T., McDonald, B., Pragidis, I., 2019. Assimilation of oil news into prices. *International Review of Financial Analysis* 63, 105–118.
- Newey, W.K., West, K.D., 1987. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica* 55, 703–708.
- Patton, A.J., Ziegel, J.F., Chen, R., 2019. Dynamic semiparametric models for expected shortfall (and value-at-risk). *Journal of Econometrics* 211, 388–413.
- Sévi, B., 2014. Forecasting the volatility of crude oil futures using intraday data. *European Journal of Operational Research* 235, 643–659.
- Shiller, R.J., 2017. Narrative economics. *American Economic Review* 107, 967–1004.
- Tauchen, G.E., Pitts, M., 1983. The price variability-volume relationship on speculative markets. *Econometrica* 51, 485–505.
- Tetlock, P.C., 2007. Giving content to investor sentiment: The role of media in the stock market. *Journal of Finance* 62, 1139–1168.
- Wei, J., Lima, L.R., Godeiro, L.L., 2026. Factor-based quantile forecasting with textual data. *Journal of Applied Econometrics* .