

Online Appendix

Forecasting Oil-Price Tail Risk with Newspaper Attention

Gabriel Abeyie

University of North Carolina Wilmington

Contents

A	Data and Text-Measure Construction	2
A.1	Text-Measure Construction Protocol	2
A.2	Robustness to the Attention Transformation	5
A.3	Sensitivity to the Oil-Window Size	6
A.4	Newspaper-Level Heterogeneity	7
A.5	Normalized Risk-Intensity Measures	7
B	Robustness of the Forecasting Results	8
B.1	Robustness of Out-of-Sample Forecasting Gains	8
B.2	Subsample Stability and Time Variation	9
B.3	Controlling for Contemporaneous Oil Returns	10
B.4	Controlling for Implied Volatility	10
B.5	Incremental Information beyond Uncertainty and Geopolitical-Risk Indices	12
C	Additional Results and Economic Relevance	15
C.1	Forecast Horse Races and Forecast Combinations	15
C.2	Economic Value of the Forecasts	16
C.3	Upside versus Downside Realized Volatility	18
C.4	Narrative-Type Decomposition	19

A Data and Text-Measure Construction

This section describes the construction of the newspaper corpus and the text measures. We first document the corpus, text preprocessing, oil-specific windows, dictionary development, narrative classification, and inclusion criteria. We then assess the sensitivity of the measures to the attention transformation, the oil-window length, the composition of the newspapers, and the normalization of risk language.

A.1 Text-Measure Construction Protocol

This subsection provides the detail needed to audit and replicate the corpus construction and the text measures.

A.1.1 Corpus construction

The corpus is constructed in ProQuest TDM Studio from *Wall Street Journal*, *New York Times*, *Washington Post*, and *Los Angeles Times*, January 1988 through December 2024. We restrict the source type to newspapers and the document type to news, feature, general, and cover-page articles. Editorials, op-eds, letters to the editor, advertisements, obituaries, and corrections are excluded from the baseline search. The baseline search is

```
("crude oil" OR "oil price" OR "oil prices" OR OPEC OR "Brent crude" OR WTI OR "West  
Texas Intermediate")  
AND  
(price* OR market* OR supply OR demand OR inventory OR inventories OR production  
OR output OR embargo* OR sanction* OR war OR conflict* OR hurricane* OR refinery  
OR refineries OR "Strategic Petroleum Reserve" OR SPR)
```

The raw export comprises 84,212 records. We remove records with missing dates or text, duplicate identifiers, and duplicate document text, together with extremely short documents; duplicates are identified first by ProQuest identifier and then by title–date–publication–text matches. The cleaned corpus contains 79,330 documents. The robustness check uses a stricter definition of oil-news coverage: it drops recurring market-digest and brief-summary items, requires an oil-related

headline or repeated oil mentions in the body, and imposes a minimum length, leaving 35,455 documents.

A.1.2 Preprocessing and oil-related windows

The text is processed minimally so that word order and phrasing are preserved: it is lower-cased, HTML tags are removed, and extra whitespace is stripped, but stop words are kept, and no stemming is applied, since the measures depend on the context around oil mentions. Oil and oil-market mentions are identified by exact matching of strings and phrases. The basic list of oil-related keywords is:

crude oil, oil price, oil prices, oil market, oil markets, petroleum,
Brent crude, Brent, WTI, West Texas Intermediate, OPEC, refinery, refineries,
gasoline, diesel, crude stocks, inventories, stockpiles, production, output, supply.

From each mention we extract a symmetric window of ± 20 words. The risk, uncertainty, and tone variables are computed within these windows; where windows overlap within an article, the overlap is counted once in the word-denominator variables.

A.1.3 Dictionaries

Oil attention is the monthly count of oil-focused articles, $Attention_t = \log(1 + N_t)$. Risk attention and oil uncertainty are both derived from a dictionary of risk- and uncertainty-related words occurring within oil-centered windows but are normalized differently. The dictionary contains

uncertain, uncertainty, unclear, volatile, volatility, risk, risks, risky,
concern, concerns, fear, fears, worried, instability, unstable, unpredictable,
doubt, speculation, tension, threat, possible, may, might, could, expected,
unexpected, surprise.

Let H_t be the monthly count of the above words within oil-centered windows and W_t be the total number of words in oil-centered windows. Risk attention is the log count, $RiskAttention_t = \log(1 + H_t)$, measuring the *salience* (i.e., the amount) of risk- and uncertainty-related words within

oil coverage. Oil uncertainty is the normalized count, $Uncertainty_t = H_t/(W_t + 1)$, measuring the *intensity* (i.e., concentration) of such words per word. Since both indicators are based on the same dictionary, they are different transformations of the same underlying signal, with the former capturing the quantity and the latter the intensity of risk- and uncertainty-related words; the empirical analysis reflects this relation.

The oil-price-pressure tone score counts upward- and downward-pressure language separately. Upward-pressure terms include

production cuts, output cuts, supply cuts, supply disruptions, shortages,
sanctions, embargoes, wars, hurricanes, outages, refinery shutdowns,
strong demand, rising demand, inventory draws, stockpile draws.

The downward pressure terms include

glut, oversupply, surplus, inventory builds, stockpile builds,
production increases, output increases, rising supply, weak demand,
demand slowdown, recession, slowing growth.

Terms that are sensitive to modifiers (e.g., production, output, inventories, stockpiles, supply, and demand) take the sign from adjacent words: “production cut,” “output cut,” “supply disruption” and “shortage” are considered to be upward price pressure terms, “production increase,” “output increase,” “rising output,” “rising supply” and “weak demand” are considered to be downward price pressure terms, “inventory draw” is considered an upward price pressure term while “inventory build” is considered a downward price pressure term. The tone score index is $Tone_t = (UpPressure_t - DownPressure_t)/(TotalWords_t + 1)$.

A.1.4 Narrative-type categories

For the decomposition in Appendix C.4, oil-centered text is classified into six categories:

- *Geopolitical/sanctions/war*: war, conflict, sanctions, embargo, attack, attacks, tensions, Iran, Iraq, Kuwait, Russia, Ukraine, Venezuela, Middle East.

- *OPEC/coordination*: OPEC, quota, quotas, production cut, output cut, supply agreement, Saudi, spare capacity, cartel.
- *Inventories/fundamentals*: inventories, inventory, crude stocks, stockpiles, EIA, API, refinery utilization, supply-demand balance, surplus, deficit.
- *Demand/macro*: recession, slowdown, weak demand, strong demand, China demand, global growth, consumption, economic growth.
- *Refining/bottlenecks*: refinery, refineries, outage, shutdown, hurricane, gasoline, diesel, crack spread, Gulf Coast.
- *Energy transition/policy*: climate policy, renewables, electric vehicles, fossil fuels, regulation, emissions, energy transition.

These indices are descriptive and the categories are not orthogonal.

A.1.5 Examples included and excluded

The baseline set contains articles that carry economically relevant information about the oil market: OPEC production cuts, crude-inventory data, and sanctions affecting oil exports. The strict set excludes articles where oil is mentioned only incidentally, routine market digests with no substantive oil content, short snippets, stock-market summaries that merely mention oil, commodity-price tables, and articles where “oil” appears only in a company name.

A.2 Robustness to the Attention Transformation

The baseline attention measure uses a logarithmic transformation, $Attention_t = \log(1 + N_t)$, which dampens the influence of outlier months. To check that the results are not an artifact of this transformation, Table A.1 reports out-of-sample improvements under four transformations of the monthly oil-article count: the baseline logarithm, the raw count, the inverse hyperbolic sine, and the percentile rank. The out-of-sample R^2 for downside realized volatility is essentially identical across the log, inverse-hyperbolic-sine, and percentile-rank transformations, and remains positive and significant for the raw count, which tempers spikes less.

Table A.1: Attention Transformation: Out-of-Sample Gain on Downside Realized Volatility

Transformation of N_t	WTI Downside RV	Brent Downside RV
$\log(1 + N_t)$ (baseline)	+3.7*** (3.81)	+4.7*** (4.21)
Raw count N_t	+2.3*** (3.39)	+2.9*** (3.50)
$\text{asinh}(N_t)$	+3.7*** (3.80)	+4.7*** (4.20)
Percentile rank	+3.8*** (4.12)	+4.7*** (4.41)

Notes: Out-of-sample R_{OOS}^2 (percentage points) from adding the transformed attention measure to a HAR benchmark, with [Clark and West \(2007\)](#) MSPE-adjusted t -statistics in parentheses. N_t is the monthly count of oil-focused articles. Each transformation is standardized before entering the forecasting model; because the model includes an intercept, standardization does not affect the forecasts. Expanding-window estimation, $T = 299$. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

A.3 Sensitivity to the Oil-Window Size

Dictionary hits for the text scores are counted within a symmetric window of ± 20 tokens around each oil mention. Two points bear on sensitivity to this choice. First, oil-news attention relies on article counts and is independent of the window, so the headline result is insensitive to the window by construction. Second, the window-based risk and uncertainty measures are insensitive to widening or narrowing the window to ± 10 or ± 30 tokens (Table A.2). Their out-of-sample gains remain positive and significant at every window width. Each window-based series is correlated above 0.97 across the ± 10 , ± 20 , and ± 30 specifications, so the gains are nearly invariant to the choice of window.

Table A.2: Window-Size Sensitivity: Out-of-Sample Gain on Downside Realized Volatility

	± 10	± 20	± 30
<i>Risk attention</i>			
WTI	+4.6*** (4.45)	+5.1*** (4.51)	+5.4*** (4.58)
Brent	+5.8*** (4.92)	+6.3*** (4.98)	+6.3*** (4.98)
<i>Uncertainty</i>			
WTI	+1.1*** (2.61)	+1.5*** (2.68)	+2.1*** (3.22)
Brent	+1.8*** (3.15)	+2.3*** (3.14)	+2.3*** (3.14)

Notes: Out-of-sample R_{OOS}^2 (percentage points) for downside realized volatility from adding each window-based measure to a HAR benchmark, with [Clark and West \(2007\)](#) t -statistics in parentheses, as the oil-context window is varied from ± 10 to ± 30 tokens. Oil-news attention is an article-count measure and does not depend on the window. Expanding-window estimation, $T = 299$. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

A.4 Newspaper-Level Heterogeneity

The baseline corpus combines four newspapers with different audiences, degrees of financial specialization, and regional focus. Table A.3 constructs a separate attention measure from each newspaper and evaluates its predictive power individually. The signal is present in all four, but its strength varies: it is strongest for the New York Times (+5.4/ + 6.8 percentage points for WTI and Brent downside RV) and the Wall Street Journal (+2.8/ + 3.8), and weaker for the Washington Post (+1.3/ + 2.0) and the Los Angeles Times (+0.7/ + 0.9, only marginally significant). The four measures are positively correlated (for example, 0.79 between the Wall Street Journal and the New York Times), pointing to a common national oil-attention signal, with the Washington Post the most idiosyncratic.

Table A.3: Per-Newspaper Attention: Out-of-Sample Gain on Downside Realized Volatility

Newspaper	WTI Downside RV	Brent Downside RV	Months with coverage
New York Times	+5.4*** (4.47)	+6.8*** (4.83)	442
Wall Street Journal	+2.8*** (3.33)	+3.8*** (3.88)	442
Washington Post	+1.3** (2.44)	+2.0*** (3.17)	331
Los Angeles Times	+0.7* (1.88)	+0.9** (2.13)	418

Notes: Out-of-sample R_{OOS}^2 (percentage points) for downside realized volatility from adding a single-newspaper attention measure to a HAR benchmark, with [Clark and West \(2007\)](#) t -statistics in parentheses. Each measure is $\log(1 + N_t^{(j)})$ for the monthly count of oil-focused articles in newspaper j . “Months with coverage” is the number of sample months in which the newspaper has at least one oil-focused article. Expanding-window estimation, $T = 299$. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

A.5 Normalized Risk-Intensity Measures

Risk attention is a salience count, so a rise in it could reflect either an expansion of oil coverage or a greater intensity of risk language within that coverage. To check that the downside-volatility relation is not driven by coverage volume alone, Table A.4 reports three normalized measures, dividing the count of risk-language terms by the number of oil-focused articles, by the number of oil-query articles, and by the number of words in the oil windows. All three still predict downside realized volatility under both benchmarks, so the relation is not purely a volume effect. This is consistent with the decomposition in Section 4.3: there is a modest risk-language content effect, but it is small and disappears once implied volatility is controlled for.

Table A.4: Normalized Risk-Intensity Measures and Downside Realized Volatility

	WTI Downside RV	Brent Downside RV
Risk hits / oil-focused article	0.070** (2.06)	0.072* (1.86)
Δ Adj. R^2	0.003	0.003
Risk hits / oil-query article	0.101*** (2.80)	0.099** (2.55)
Δ Adj. R^2	0.008	0.008
Risk hits / 10,000 oil-window words	0.068** (2.07)	0.092*** (2.67)
Δ Adj. R^2	0.003	0.007
Benchmark Adj. R^2	0.200	0.169
Observations	442	442

Notes: Predictive regressions of next-month downside realized volatility on current downside volatility and one normalized risk-attention measure. Newey and West (1987) t -statistics in parentheses. Δ Adj. R^2 is relative to the benchmark with current downside volatility only. $T = 442$. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

B Robustness of the Forecasting Results

This section examines the robustness of the paper’s main out-of-sample forecasting results. We first consider alternative evaluation periods, exclude the COVID-19 oil-market disruption, and tighten the definition of the oil-news corpus (Section B.1). We then assess stability across subsamples and over the transition from print to digital newspapers (Section B.2). Finally, we ask whether newspaper attention carries predictive information beyond oil returns, option-implied volatility, and the Economic Policy Uncertainty and Geopolitical Risk indices (Sections B.3–B.5). Taken together, the results show that the forecasting gains cannot be attributed to a particular sample period, corpus, or existing risk measure.

B.1 Robustness of Out-of-Sample Forecasting Gains

Table B.1 reports the out-of-sample R_{OOS}^2 for the attention-augmented model under three checks. In Panel A we shift the start of the evaluation period across five dates between 1998 and 2010, which changes both the initial estimation sample and the evaluation window. In Panel B we drop the COVID crash months, the most volatile episode in the sample. Panel C uses the strict corpus, which excludes market digests and low-relevance articles. In all three, oil-news coverage continues to improve forecasts of realized volatility, and of downside realized volatility in particular. Every

improvement is significant by the [Clark and West \(2007\)](#) test.

Table B.1: Robustness of Out-of-Sample Forecasting Gains

Robustness check	Benchmark	N	WTI		Brent	
			RV	Downside RV	RV	Downside RV

Panel A: Alternative out-of-sample start dates

Start 1998	AR	323	1.86***	4.44***	1.82***	4.74***
	HAR	323	2.20***	4.67***	2.79***	5.60***
Start 2000	AR	299	2.18***	4.87***	2.23***	5.48***
	HAR	299	2.46***	5.08***	3.23***	6.30***
Start 2003	AR	263	2.42***	4.77***	2.76***	5.89***
	HAR	263	2.71***	4.94***	3.80***	6.32***
Start 2005	AR	239	2.53***	5.29***	3.06***	6.42***
	HAR	239	2.79***	5.26***	4.00***	6.88***
Start 2010	AR	179	3.21***	6.93***	3.43***	7.46***
	HAR	179	3.18***	6.21***	4.18***	7.36***

Panel B: Excluding COVID oil-crash months

Excl. COVID	AR	295	2.94***	5.92***	3.06***	7.08***
	HAR	295	3.46***	6.17***	4.56***	7.99***

Panel C: Strict oil-news corpus

Strict corpus	AR	299	1.61***	4.42***	1.93***	5.28***
	HAR	299	2.01***	5.18***	2.97***	6.54***

Notes: Out-of-sample R^2_{OOS} (percentage points) for the attention-augmented model relative to the indicated benchmark. Significance uses the [Clark and West \(2007\)](#) MSPE-adjusted test for nested forecasts. Panel A varies the evaluation start date; Panel B excludes target months between March and June 2020; Panel C uses the strict corpus. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

B.2 Subsample Stability and Time Variation

Newspaper circulation has fallen since the mid-2000s as the media environment shifted to digital sources, raising the question of whether the predictive content is an artifact of the print era. Table [B.2](#) splits the out-of-sample period into a print-dominant period (2000–2010) and a digital-dominant period (2011–2024). Downside-volatility gains are significant in both eras and for both crudes; if anything, they are larger in the digital era, so the predictability is not an artifact of the print period. The bottom panel reports gains over four sub-periods. Predictability varies over time in an interpretable way: low in the calm early 2000s, high over 2006–2017, which spans the 2008 spike and the 2014–2016 decline, and moderate again over 2018–2024. The downside-volatility gain is positive in about 90% of five-year windows.

Table B.2: Subsample Stability: Out-of-Sample Gain from Attention vs. HAR

Panel A: Print- and digital-dominant periods

	Full sample 2000–2024	Print-dominant 2000–2010	Digital-dominant 2011–2024
WTI RV	1.9***	0.8	2.4***
WTI Downside RV	3.7***	3.1**	4.1***
Brent RV	2.9***	1.2	3.7***
Brent Downside RV	4.7***	4.4**	4.9***

Panel B: Downside RV by sub-period

	2000–2005	2006–2011	2012–2017	2018–2024
WTI Downside RV	1.2	6.1**	9.8**	1.0*
Brent Downside RV	1.3	9.5***	15.0***	1.0**

Notes: Out-of-sample R_{OOS}^2 (percentage points) from adding oil-news attention to a HAR benchmark, evaluated over the indicated subsamples of the 2000:01–2024:11 forecast period, with significance from the [Clark and West \(2007\)](#) test. Forecasts are generated recursively on the full history; the subsamples restrict only the evaluation window. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

B.3 Controlling for Contemporaneous Oil Returns

If oil coverage were driven by the current month’s price change, the narrative coefficients could pick up an existing return–volatility relation rather than independent information in the news. [Table B.3](#) adds the contemporaneous oil return and its absolute value to the baseline regressions, controlling for the current month’s price change in both sign and magnitude. The attention measures remain significant at the 1% level for all four outcomes. The coefficients fall slightly; the largest drop is for risk attention on Brent downside volatility, from 0.213 to 0.162, but they remain economically meaningful and well estimated. Oil coverage therefore carries information about next-month volatility that is not contained in current price movements.

B.4 Controlling for Implied Volatility

If newspaper attention only reproduced information already contained in option-implied volatility, it would add little to a benchmark that includes an implied-volatility index ([Bekaert and Hoerova, 2014](#)). We use two controls: the VIX, the option-implied measure of expected S&P 500 volatility, available from January 1990; and the OVX, the oil-specific implied-volatility index, available from June 2007. The unconditional correlation between oil coverage and the log VIX is about 0.15, with a similarly modest correlation to the log OVX, so the overlap is small even before formal controls.

Table B.3: Predictive Regressions Controlling for Contemporaneous Oil Returns

Narrative measure	WTI		Brent	
	RV	Downside RV	RV	Downside RV
Oil Attention	0.088*** (2.58)	0.152*** (3.16)	0.091*** (2.84)	0.143*** (3.41)
Δ Adj. R^2	0.007	0.018	0.008	0.015
Oil Risk Attention	0.093*** (2.81)	0.169*** (3.60)	0.088*** (2.89)	0.162*** (4.04)
Δ Adj. R^2	0.008	0.022	0.007	0.020
Oil Uncertainty	0.028 (1.14)	0.066* (1.91)	0.030 (1.23)	0.091** (2.43)
Δ Adj. R^2	−0.000	0.003	0.000	0.007
Oil Tone	−0.027 (−0.69)	0.002 (0.04)	−0.013 (−0.37)	0.035 (0.78)
Δ Adj. R^2	−0.000	−0.002	−0.001	−0.000
Benchmark Adj. R^2	0.420	0.236	0.477	0.267
Observations	442	442	442	442

Notes: Predictive regressions of one-step-ahead volatility on current volatility, the contemporaneous monthly oil return, its absolute value, and one standardized narrative measure. [Newey and West \(1987\)](#) t -statistics are reported in parentheses. Δ Adj. R^2 is measured relative to the benchmark containing current volatility and the two return controls. $T = 442$. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B.4 adds each implied-volatility index to the predictive regressions. With the VIX (Panel A), oil attention and risk attention stay significant at the 1% level across all four outcomes, and the VIX itself is significant throughout. With the OVX over the shorter post-2007 sample (Panel B), the attention measures again hold at the 1% level. Oil uncertainty, marginal in the VIX panel, strengthens under the OVX control, while oil tone stays indistinguishable from zero.

The out-of-sample test is repeated in Table B.5. Adding oil coverage to the AR+VIX benchmark improves the R_{OOS}^2 by 3.4 to 6.4 percentage points, and the HAR+VIX benchmark by 1.8 to 5.7 points. In the shorter OVX sample, the gains range from 3.6 to 5.5 points against AR+OVX and from 5.4 to 9.4 points against HAR+OVX. All sixteen cases are significant at the 1% level, and in every panel the downside gains exceed the total-volatility gains. Attention thus cannot be regarded as a newspaper-based analog of either implied-volatility index.

These results show that attention carries predictive information beyond the implied-volatility benchmarks. The reverse question, whether the risk dictionary adds anything once implied volatility is accounted for, is the one taken up in Section 4.3. Table B.6 runs the encompassing test. Each entry is the R_{OOS}^2 from adding risk attention to a benchmark that already contains both the implied-volatility index and plain attention; a positive, significant value would indicate that risk weighting matters beyond coverage volume and option-implied volatility. None is significant. Against volatility-history benchmarks the same test produces downside gains of 1.1 and 1.3 points (Table 6, Panel B), but conditioning on implied volatility makes them vanish. The reverse, adding plain attention to a benchmark that contains risk attention, is also insignificant in every case.

B.5 Incremental Information beyond Uncertainty and Geopolitical-Risk Indices

The previous subsection (Appendix B.4) showed that oil-news attention survives controls for option-implied volatility. Here we extend the same test to the two most widely used text-based risk indices: the Economic Policy Uncertainty index of Baker et al. (2016) and the Geopolitical Risk index of Caldara and Iacoviello (2022). Attention is only weakly correlated with either (-0.17 with the log EPU, $+0.17$ with the log GPR), and Table B.7 reports its incremental content. In-sample (Panel A), attention remains significant at the 1% level for all four outcomes after controlling for both indices, while the indices themselves do not forecast oil volatility once current volatility is included. Out of sample (Panel B), adding attention improves forecasts of downside realized volatility by

Table B.4: Predictive Regressions Controlling for Implied Volatility

Narrative measure	WTI		Brent	
	RV	Downside RV	RV	Downside RV
<i>Panel A: Augmented with the log VIX ($T = 419$)</i>				
Oil Attention	0.128*** (3.14)	0.203*** (3.71)	0.141*** (3.74)	0.236*** (4.64)
Δ Adj. R^2	0.016	0.034	0.020	0.045
Oil Risk Attention	0.119*** (3.36)	0.202*** (4.18)	0.121*** (3.58)	0.229*** (5.05)
Δ Adj. R^2	0.014	0.034	0.015	0.044
Oil Uncertainty	0.029 (1.12)	0.065* (1.88)	0.023 (0.89)	0.080** (2.16)
Δ Adj. R^2	-0.000	0.002	-0.001	0.005
Oil Tone	-0.022 (-0.66)	0.009 (0.21)	-0.021 (-0.69)	0.018 (0.45)
Δ Adj. R^2	-0.001	-0.002	-0.001	-0.002
<i>Panel B: Augmented with the log OVX ($T = 210$)</i>				
Oil Attention	0.196*** (3.58)	0.270*** (3.31)	0.200*** (3.93)	0.277*** (3.54)
Δ Adj. R^2	0.027	0.040	0.028	0.042
Oil Risk Attention	0.197*** (3.95)	0.281*** (3.68)	0.180*** (3.94)	0.279*** (3.85)
Δ Adj. R^2	0.028	0.046	0.024	0.045
Oil Uncertainty	0.108*** (2.75)	0.152*** (2.71)	0.088** (2.19)	0.160*** (2.99)
Δ Adj. R^2	0.010	0.015	0.006	0.017
Oil Tone	-0.009 (-0.21)	0.026 (0.46)	-0.007 (-0.19)	0.037 (0.69)
Δ Adj. R^2	-0.002	-0.003	-0.002	-0.002
Panel A benchmark Adj. R^2	0.425	0.235	0.461	0.208
Panel B benchmark Adj. R^2	0.479	0.294	0.543	0.301

Notes: Predictive regressions of one-step-ahead volatility on current volatility, a log implied-volatility index, and one standardized narrative measure. Panel A uses the log VIX; Panel B the log OVX. [Newey and West \(1987\)](#) t -statistics in parentheses. Δ Adj. R^2 is relative to the benchmark with current volatility and the implied-volatility index. The VIX is available from January 1990 ($T = 419$), the OVX from June 2007 ($T = 210$). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B.5: Out-of-Sample Forecasting Gains Controlling for Implied Volatility

Narrative augmentation	WTI		Brent	
	RV	Downside RV	RV	Downside RV
<i>Panel A: AR + VIX benchmark ($N = 299$)</i>				
Oil Risk Attention	3.44*** (3.91)	5.73*** (4.40)	3.47*** (3.43)	6.44*** (4.21)
<i>Panel B: HAR + VIX benchmark ($N = 299$)</i>				
Oil Risk Attention	1.75*** (2.65)	4.39*** (4.16)	2.51*** (2.91)	5.72*** (4.59)
<i>Panel C: AR + OVX benchmark ($N = 179$)</i>				
Oil Risk Attention	3.61*** (3.76)	4.66*** (3.86)	3.67*** (3.34)	5.51*** (4.20)
<i>Panel D: HAR + OVX benchmark ($N = 175$)</i>				
Oil Risk Attention	5.40*** (3.79)	8.36*** (4.85)	7.05*** (3.62)	9.41*** (4.63)

Notes: Out-of-sample R^2_{OOS} (percentage points) from adding oil coverage to benchmarks that already control for implied volatility. Panels A–B use the VIX; Panels C–D the OVX. Significance uses the [Clark and West \(2007\)](#) MSPE-adjusted test. The VIX panels use 2000:01–2024:11 ($N = 299$); the OVX panels use the shorter post-OVX sample, with AR+OVX from 2010:01 ($N = 179$) and HAR+OVX from 2010:05 ($N = 175$), the later HAR start reflecting the trailing averages needed within the OVX sample. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B.6: Encompassing Test: Does Risk Attention Add Beyond Attention and Implied Volatility?

Benchmark (contains attention)	WTI		Brent	
	RV	Downside RV	RV	Downside RV
AR + VIX	−0.3 (−1.07)	−0.2 (0.02)	−0.4 (−0.37)	−0.3 (−0.20)
HAR + VIX	−0.2 (−1.34)	+0.2 (1.00)	−0.4 (−1.79)	+0.2 (0.87)
AR + OVX	−0.6 (0.23)	−0.3 (0.44)	−0.9 (0.34)	−0.3 (0.64)
HAR + OVX	−1.3 (−1.15)	−0.3 (0.24)	−1.1 (−0.71)	−0.2 (0.15)

Notes: Each entry is the out-of-sample R_{OOS}^2 (percentage points) from adding risk attention to a benchmark that already contains the indicated implied-volatility index and plain oil-news attention, with [Clark and West \(2007\)](#) MSPE-adjusted t -statistics in parentheses. A positive, significant value would indicate that risk-weighted coverage forecasts beyond plain coverage and option-implied volatility. None is significant at the 5% level. The reverse test, adding plain attention to a model containing risk attention, is also insignificant throughout (not reported). For comparison, against volatility-history benchmarks without implied volatility the downside gains are +1.1 (WTI) and +1.3 (Brent) and significant (Table 6). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

2.9 to 4.3 percentage points against benchmarks that already contain EPU and GPR, and by 3.3 (WTI) and 5.3 (Brent) against a benchmark that contains current and recent realized volatility, the VIX, the OVX, EPU, and GPR simultaneously. Whatever oil-news attention captures is not spanned by these measures, even jointly.

C Additional Results and Economic Relevance

This section complements the baseline forecasting evidence and examines its economic relevance. Appendix C.1 compares attention with other predictors and with forecast combinations. Appendix C.2 assesses the economic value of the forecasts. Appendix C.3 tests whether the results are specific to downside rather than upside volatility, and Appendix C.4 breaks aggregate attention into types of oil-market narrative. Together these clarify what drives the predictive gains and why they matter economically.

C.1 Forecast Horse Races and Forecast Combinations

Table C.1 compares attention with alternative text measures and with traditional uncertainty and volatility indices in the same forecasting framework. Panel A adds each predictor separately to the HAR benchmark. The text-based attention measures deliver the largest improvements for downside realized volatility: risk attention and aggregate attention give the biggest gains, followed by oil uncertainty. The VIX, EPU, and GPR indices, by contrast, do not improve on the HAR

Table B.7: Oil-News Attention beyond EPU and GPR

	WTI		Brent	
	RV	Downside RV	RV	Downside RV
<i>Panel A: In-sample attention coefficient, controlling for log EPU and log GPR</i>				
Oil Attention	0.103*** (2.68)	0.190*** (3.53)	0.115*** (3.02)	0.213*** (4.09)
Δ Adj. R^2	0.009	0.028	0.012	0.034
<i>Panel B: Out-of-sample R_{OOS}^2 (pp) from adding attention</i>				
AR + EPU + GPR	+1.7** (2.42)	+3.6*** (3.38)	+2.4*** (2.67)	+4.3*** (3.20)
HAR + EPU + GPR	+1.4** (2.37)	+2.9*** (3.26)	+2.6*** (3.03)	+3.9*** (3.33)
HAR + VIX + OVX + EPU + GPR (post-2010, $N = 179$)	+2.1*** (3.17)	+3.3*** (3.78)	+5.5*** (4.30)	+5.3*** (4.20)

Notes: Panel A reports the coefficient on oil-news attention in a predictive regression that also includes current realized volatility, the log EPU index, and the log GPR index, with [Newey and West \(1987\)](#) t -statistics (six lags) in parentheses; Δ Adj. R^2 is relative to the same regression without attention. Panel B reports out-of-sample R_{OOS}^2 (percentage points) from adding attention to each benchmark, with [Clark and West \(2007\)](#) MSPE-adjusted t -statistics in parentheses. The first two benchmarks use the full 2000:01–2024:11 sample; the last uses the post-OVX sample from 2010. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

benchmark, and the directional tone measure worsens the forecast. Panel B turns to forecast combinations. An equal-weight combination of the traditional predictors improves modestly on the HAR benchmark, and adding oil-news attention improves the combined forecast substantially. The attention-only model dominates the unrestricted combination.

C.2 Economic Value of the Forecasts

Beyond the Value-at-Risk results in the main text, this subsection asks whether the statistical gains translate into economic value in a volatility-timing exercise. A mean-variance investor scales oil exposure by the ratio of a target volatility to the one-month-ahead volatility forecast, with the exposure bounded between zero and two. We compare exposure timed on the attention-augmented forecast with exposure timed on the HAR forecast. Table C.2 reports the differences in annualized Sharpe ratio and certainty-equivalent return. In every case the Sharpe-ratio improvement is positive, and larger when exposure is scaled by the downside-volatility forecast, consistent with the paper’s downside focus. The certainty-equivalent improvement is swamped by the 2020 crash in the full sample but turns positive once 2020 is excluded, again largest for downside-timed exposure.

Table C.1: Forecast Horse Races and Forecast Combinations: Downside Realized Volatility

	WTI Downside RV	Brent Downside RV
<i>Panel A: R_{OOS}^2 (pp) from adding one predictor to HAR</i>		
Risk attention	+5.1*** (4.51)	+6.3*** (4.98)
Oil attention	+3.7*** (3.81)	+4.7*** (4.21)
Uncertainty	+1.5*** (2.68)	+2.3*** (3.14)
VIX	+0.8** (2.00)	−1.1 (1.62)
GPR	−0.4 (−0.76)	−0.2 (−0.06)
EPU	−0.7 (−0.70)	−0.9 (−0.69)
Tone	−1.7 (−0.81)	−1.3 (−1.33)
<i>Panel B: Equal-weight forecast combination (R_{OOS}^2, pp)</i>		
Combination of VIX, EPU, GPR	+0.9	+0.5
plus oil attention	+2.3	+2.6
Oil attention alone	+3.7	+4.7
DM statistic, gain from adding attention	2.99***	3.48***

Notes: Panel A reports out-of-sample R_{OOS}^2 (percentage points) from adding one predictor at a time to a HAR benchmark, with [Clark and West \(2007\)](#) t -statistics in parentheses. Panel B reports the R_{OOS}^2 of equal-weight combinations of the indicated single-predictor HAR forecasts, and the [Diebold and Mariano \(1995\)](#) statistic for the reduction in loss from adding the attention-augmented forecast to the combination of standard predictors. Expanding-window estimation, $T = 299$. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table C.2: Economic Value: Volatility-Timed Oil Exposure, Attention vs HAR

Asset	Sample	Total-vol timed		Downside-vol timed	
		Δ Sharpe	Δ CEQ (pp)	Δ Sharpe	Δ CEQ (pp)
WTI	Full	+0.011	−0.17	+0.016	−0.75
WTI	Ex-2020	+0.009	−0.09	+0.016	+0.10
Brent	Full	+0.010	−0.40	+0.020	−1.42
Brent	Ex-2020	+0.011	+0.13	+0.025	+1.24

Notes: Difference in annualized Sharpe ratio and annualized certainty-equivalent return (percentage points, risk aversion $\gamma = 3$) between oil exposure timed on the attention-augmented volatility forecast and exposure timed on the HAR forecast. Exposure is set to a target volatility divided by the volatility forecast, bounded between zero and two. “Total-vol timed” scales by the total realized-volatility forecast; “Downside-vol timed” scales by the downside realized-volatility forecast. Ex-2020 excludes calendar-year 2020. The out-of-sample period is 2000:01–2024:11.

C.3 Upside versus Downside Realized Volatility

The downside emphasis is a matter of degree. Several oil-market stress narratives, including wars, sanctions, and supply disruptions, can drive large positive moves rather than negative ones, so it is worth asking whether the signal is specific to the downside or reflects broader tail risk. We construct upside realized volatility from positive daily returns only,

$$URV_t = \sum_{d \in t} r_d^2 \mathbf{1}\{r_d > 0\},$$

and re-estimate the baseline regressions with $\log(1 + URV_{t+1})$. Table C.3 shows that attention and risk attention remain significant at the 1% level for both crudes, so coverage forecasts large positive moves too. The coefficients are smaller than on the downside, however: risk attention is 0.109 for WTI upside volatility against 0.193 for downside, and 0.102 versus 0.213 for Brent. The signal is present on both sides and strongest on the downside, which is the tail-risk reading the paper adopts.

Table C.3: Predictive Regressions of Upside Realized Volatility on Oil Narrative Indices

Narrative measure	WTI Upside RV	Brent Upside RV
Oil Attention	0.1129*** (3.09)	0.1227*** (3.29)
Δ Adj. R^2	0.012	0.015
Oil Risk Attention	0.1087*** (3.20)	0.1022*** (2.92)
Δ Adj. R^2	0.010	0.009
Oil Uncertainty	0.0044 (0.14)	-0.0286 (-0.95)
Δ Adj. R^2	-0.002	-0.001
Oil Tone	-0.0462 (-1.31)	-0.0466 (-1.40)
Δ Adj. R^2	0.001	0.001
Baseline AR Adj. R^2	0.312	0.321
Observations	442	442

Notes: Predictive regressions of one-step-ahead upside realized volatility, $\log(1 + URV_{t+1})$, on current upside volatility and one standardized narrative measure. Newey and West (1987) t -statistics in parentheses. Δ Adj. R^2 is relative to the AR-only baseline. $T = 442$. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

C.4 Narrative-Type Decomposition

It is natural to ask which types of oil-related narrative carry the signal in the aggregate measure. We classify oil-centered text into six topic groups and re-estimate the baseline regression with one group at a time. Table C.4 reports the results, ordered by each group’s mean incremental adjusted R^2 . The demand and macro group adds the most, followed by the geopolitical and the inventory/fundamentals groups, which share the aggregate measure’s downside tilt. The remaining three groups contribute less.

Table C.4: Predictive Regressions by Oil Narrative Type

Narrative measure	WTI		Brent	
	RV	Downside RV	RV	Downside RV
<i>Aggregate index</i>				
Oil Attention	0.093*** (2.72)	0.176*** (3.57)	0.103*** (3.08)	0.197*** (4.15)
Δ Adj. R^2	0.008	0.026	0.011	0.046
<i>Narrative-type sub-indices, ordered by mean Δ Adj. R^2</i>				
Demand / Macro	0.132*** (3.62)	0.226*** (4.26)	0.133*** (3.91)	0.254*** (4.80)
Δ Adj. R^2	0.016	0.041	0.016	0.051
Geopolitical / Sanctions / War	0.095*** (2.67)	0.192*** (3.86)	0.106*** (3.10)	0.229*** (4.79)
Δ Adj. R^2	0.008	0.031	0.010	0.042
Inventories / Fundamentals	0.095*** (2.76)	0.179*** (3.48)	0.101*** (3.06)	0.213*** (4.28)
Δ Adj. R^2	0.008	0.026	0.009	0.037
OPEC / Coordination	0.070** (2.07)	0.137*** (3.06)	0.085** (2.48)	0.150*** (3.32)
Δ Adj. R^2	0.004	0.015	0.006	0.018
Refining / Bottlenecks	0.056* (1.83)	0.121*** (2.70)	0.059** (1.97)	0.132*** (3.04)
Δ Adj. R^2	0.002	0.012	0.003	0.015
Energy Transition / Policy	0.061** (2.20)	0.089** (2.25)	0.057** (2.21)	0.097** (2.39)
Δ Adj. R^2	0.003	0.006	0.003	0.008
Baseline AR Adj. R^2	0.397	0.200	0.430	0.169
Observations	442	442	442	442

Notes: Each block reports a predictive regression of one-step-ahead volatility on current volatility and one standardized narrative-type index, constructed from oil-centered windows and listed in descending order of mean Δ Adj. R^2 . The aggregate Oil Attention row reproduces the headline result for comparison. [Newey and West \(1987\)](#) t -statistics in parentheses. The categories share vocabulary and are not orthogonal, so the decomposition is descriptive. $T = 442$. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

References

- Baker, S. R., Bloom, N., and Davis, S. J. (2016). Measuring economic policy uncertainty. *Quarterly Journal of Economics*, 131(4):1593–1636.
- Bekaert, G. and Hoerova, M. (2014). The VIX, the variance premium and stock market volatility. *Journal of Econometrics*, 183(2):181–192.
- Caldara, D. and Iacoviello, M. (2022). Measuring geopolitical risk. *American Economic Review*, 112(4):1194–1225.
- Clark, T. E. and West, K. D. (2007). Approximately normal tests for equal predictive accuracy in nested models. *Journal of Econometrics*, 138(1):291–311.
- Diebold, F. X. and Mariano, R. S. (1995). Comparing predictive accuracy. *Journal of Business and Economic Statistics*, 13(3):253–263.
- Newey, W. K. and West, K. D. (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, 55(3):703–708.