



# Macro fundamentals or geopolitical events? A textual analysis of news events for crude oil<sup>☆</sup>

Michael W. Brandt<sup>a,\*</sup>, Lin Gao<sup>b</sup>

<sup>a</sup> Fuqua School of Business, Duke University, 100 Fuqua Drive, Box 9012, Durham, NC 27708, USA

<sup>b</sup> Luxembourg School of Finance, University of Luxembourg, 6 Rue Richard Coudenhove-Kalergi, 1359, Luxembourg



## ARTICLE INFO

### JEL classification:

G12

G13

### Keywords:

Crude oil

News analytics

Sentiment

Information

## ABSTRACT

News about macroeconomic fundamentals and geopolitical events affect crude oil markets differently. Using sentiment scores for a broad set of global news of different types, we find that news related to macro fundamentals have an impact on the oil price in the short run and significantly predict oil returns in the long run. Geopolitical news have a much stronger immediate impact but exhibit no predictability. Moreover, geopolitical news generate more uncertainty and greater trading volume, consistent with a disagreement explanation, while macroeconomic news are associated with subsequent lower trading volume. Finally, we find that news sentiment tracks the statistical releases quite well and can partially predict the future realizations of the economic data.

## 1. Introduction

Oil prices are hard to explain and predict. Hamilton (2008) suggests that the real price of oil follows a random walk without drift. Baumeister and Kilian (2014), who explore an exhaustive set of oil pricing factors compiled from the literature, conclude that the explanatory power of these factors vary over time and that different factors are important at different time horizons. The ever changing nature of this predictive relationship contributes to the difficulty of forecasting oil prices. Furthermore, oil prices are not only related to economic fundamentals but also to geopolitical events that are much harder to quantify. For example, policy issuance could be a one-time event for which it is impossible to set up a time series record to then quantitatively relate this event to oil prices.

News analytics provides a way to quantify both macroeconomic and geopolitical events. It not only offers timely analysis of the news content but also captures both the raw information as well as the market perception of the news. We use news analytics to compare and contrast the importance of macroeconomic and geopolitical information for crude oil.

There are different ways news could affect oil prices and trading. On one hand, news could convey information on the current important state variables, that are otherwise only imperfectly observable in real time. Also, news could serve as an update of the changes in these state variables. These factors might further shift expectations about future oil supply and demand, which could impact the inventory decisions and oil prices today (Kilian and Murphy, 2014). On the other hand, news could affect trader sentiment, including that of the noise traders.

We consider a broad cross-section of macroeconomic and geopolitical news. The goal of this paper is to investigate the role of these different types of news for both crude oil prices and trading activity in oil markets. We rely on news sentiment scores provided by RavenPack to capture the sign, magnitude, relevance, and novelty of the news. Our results highlight important differences between the roles of macroeconomic and geopolitical news. We find that macroeconomic news account for around three fourth of the total

<sup>☆</sup> We thank participants at the NBER Economics of Commodity Markets Meeting, SoFiE 2017, as well as two anonymous referees for helpful comments and suggestions.

\* Corresponding author.

E-mail addresses: [mbrandt@duke.edu](mailto:mbrandt@duke.edu) (M.W. Brandt), [heikegl@gmail.com](mailto:heikegl@gmail.com) (L. Gao).

explained variations in oil returns, whereas geopolitical news make up the rest. News about economic growth are the strongest predictors of oil returns over coming months, confirming the findings of Kilian (2009) that macroeconomic growth generates demand for oil. In contrast, the response of oil prices to geopolitical events is strong and immediate. News are incorporated into the oil price at once without follow-on momentum, consistent with informational efficiency.

Why do macroeconomic news and geopolitical news impact oil prices differently? The reason may be attributed to the different characteristics of these two groups of news. Geopolitical news such as terrorism or natural disasters often happen abruptly, and the duration of such news is usually short. Therefore, geopolitical news typically generate more uncertainty and the predictability is relatively weaker than macro news. The short duration of the news means that the impact duration is also short, because oil prices may be driven by the next influential news a few days later. In contrast, macro news are prevalent in our daily life on an on-going basis. Our expectation on the current economic state is thus updated continuously. Therefore, the economic news come less as a surprise compared to the geopolitical news.

To confirm such hypothesis, we explore the relationship between the news indices and the actual statistical release. We provide evidence that our macro news indices track the corresponding real statistical releases quite well, and can partially predict the outcome of the statistics, even after controlling for the survey data. As the news indices are available at a higher frequency such as at a daily level than those macro statistics at monthly or quarterly frequency, the news indices can provide a timely measure of the current state. In addition, for some important macro statistics such as FOMC announcement or durable goods orders, news are good at catching the downward turning point while remaining conservative in picking up the upward trend. Hence, the predictive ability of economic growth news suggests that this type of news develop more gradually. These change beliefs more smoothly over time instead of causing large uncertainty and jumps in prices. The gradual diffusion of public information and resulting momentum and predictability in oil returns may be attributed either to behavioral reasons such as investor inattention or to the interaction of different types of traders who only trade based on news or news-based price movements (Hong and Stein, 1999). However, it is beyond the scope of our paper to distinguish these cases.

We also explore the impact of news on trading volume. On one hand, Peress (2014) argues that media contributes to the efficiency of markets by improving the dissemination of information among investors and its incorporation into prices. On the other hand, sentiment theories predict that noise traders and rational traders disagree when the tone of news is extremely positive or negative. Disagreement increases trading volume because noise traders follow the media sentiment (Antweiler and Frank, 2004). We find that geopolitical news such as terrorism, war and conflict, civil unrest, and natural disasters cause a statistically significant increase in the trading volume of crude oil for the next two to three days. This evidence suggests that these geopolitical events cause more future uncertainty and investors disagree about the future oil price trajectory regarding the pending risks. In contrast, macroeconomic news such as consumer spending, durable goods orders, housing, economic growth, CPI and exports dampens subsequent trading volume. This may be because the current news come as no surprise to traders, so they interpret the news in the same way and trading is not necessary. In further examination of the net positioning of futures traders, we discover that non-commercial traders tend to move in the same direction as the news release. For example, they reduce their long position following negative geopolitical news, or increase it after positive macroeconomic news. Commercial traders react in the opposite direction.

We uncover the differences between news from the most important oil-producing versus oil-consuming countries in certain news categories. For example, civil unrest in the major oil-consuming countries signal often the economic contraction and decrease consumption. Such news are typically positively correlated with oil prices. In contrast, civil unrest in the major oil-producing countries cause concern of supply disruption and are negatively related to oil prices. Though news on geopolitics and macro economy may be important for the oil pricing, news are only very weakly related to the contemporaneous and future actual oil consumption and demand growth, suggesting that the sentiment effect dominates the risks these events actually cause.

Our paper is part of a growing body of finance and accounting research utilizing textual analysis to examine the tone and sentiment of information. Among the first studies in this field, Tetlock (2007) identifies media pessimism by employing the textual analysis program General Inquirer and the Harvard-IV-4 dictionary to count the negative words in the Abreast of the Market column of *The Wall Street Journal*. Loughran and McDonald (2011) develop an alternative, more finance focused, word list and link negative sentiment to stock returns, trading volume, and return volatility. Das and Chen (2007) employ an algorithm comprising different classifier algorithms coupled together by a voting scheme to extract investor sentiment from stock message boards. Some studies rely on the news analytics tools. For example, Dzieliński (2012) and Dzieliński and Hasseltoft (2015) rely on the Thomson Reuters News Analytics. Dang et al. (2015) utilize RavenPack News Analytics.

Most work in this literature deals either with the general equity market and aggregate news about equities and the economy (Dougal et al., 2012) or with the cross-section of firms and firm specific news such as mergers and acquisitions, product launches, or earnings announcements (see, for example, Fang and Peress (2009), Boudoukh et al. (2013) or Uhl (2014)). Soo (2015) studies an important single market like us (the housing market in her case) but she focuses on only housing related news sentiment. Similarly, Smales (2014) or Smales (2015), investigate the gold-specific news impact on gold market. Besides the unique focus on crude oil, our paper is different from the aforementioned studies in that we consider a much broader set of news categories as inputs, rather than just micro-level supply and demand factors that are likely endogenous to macroeconomic and geopolitical developments. Before we can study how news sentiment affects oil prices, we first have to document what type of news is relevant. This additional step is as an important differentiator of our paper. In a nutshell, our approach of looking at a broad set of non-oil-specific news as well as oil-specific news is more general and more interesting for investment decision making.

Our results contribute to the literature on news-related price patterns in other contexts. Tetlock (2007) finds that high pessimism predicts both downward price pressure followed by reversion to fundamentals and high trading volume. Chan (2003) finds that stocks with bad news in particular exhibit momentum, especially in the case of small, low priced and illiquid stocks due to limits to arbitrage

and slow investor reaction, while stocks without news do not. He finds that investors overreact to price shocks, causing excessive trading volume and volatility, which ultimately leads to reversals. Less directly related, [Garcia \(2013\)](#) find that the predictability of stock returns using news content is concentrated in recessions.

Working with macroeconomic news analytics is different than working with regularly scheduled macroeconomic releases. [Petersen \(2004\)](#) distinguishes between “hard” and “soft” information. He characterizes “hard” information as quantitative and easily processable by computers, whereas “soft” information is qualitative and harder to interpret by computers. Based on this classification, [Zhang \(2013\)](#) finds that soft information proxied by news analytics is incorporated into prices more slowly over time. The macroeconomic news analytics we study capture similar soft news and therefore also has the potential to exhibit slower transmission to prices. In this sense, our paper differentiates from the existing literature focusing on responses of crude oil prices to scheduled macro news announcement from U.S. ([Kilian and Vega, 2011](#)) or forecasting of crude oil prices with macro data (see for example [Baumeister and Kilian \(2012\)](#) or [Baumeister and Kilian \(2015\)](#)). Another difference from papers on the schedule announcement is that our news indices can be available at a higher frequency than the scheduled releases with a fixed frequency. In this paper, we employ daily data. In this aspect, we also differentiate from studies examining the scheduled macro announcement effect on high frequency crude oil prices ([Elder et al., 2013](#)). The reason that we focus on daily instead of high-frequency data is that the time stamp of a news release, despite of its precision in time, often does not match the intraday timing of the underlying information event ([Tetlock, 2010](#)).

Finally, our study is related to the literature on investigating different sources of oil price fluctuations.<sup>1</sup> For example, [Kilian \(2009\)](#) uses an econometric model to decompose oil price shocks into three sources: crude oil supply shocks, shocks to the global demand for all industrial commodities, and demand shocks that are specific to the global crude oil market. He finds oil-specific demand shocks cause an immediate, persistent and large increase in the real price of crude oil; an increase in the aggregate demand for all industrial commodities causes a delayed, but sustained increase; crude oil production shocks cause a small and transitory increase in oil price. Similarly, [Hitzemann \(2016\)](#) find long-run economic growth is main driver of oil future return, whereas supply and short-run economic shocks are responsible for the volatility of oil prices. [Rapport \(2013\)](#) distinguish between demand and supply driven component of crude oil returns by examining its correlation with the equity market. As with any econometric model, these findings crucially depend on the underlying model structure and assumptions. Through our news classification, we are able to corroborate these effects directly using a data-driven approach.

## 2. Data

### 2.1. News data

We measure news sentiment using the global macro package of RavenPack News Analytics Dow Jones Edition 3.0. The RavenPack global macro package tracks and analyzes information on geographical entities, government organizations, all major currencies and traded commodities. The underlying information sources include the *Dow Jones Financial Wire*, *Barron's*, and *The Wall Street Journal*. For each identified entity, RavenPack covers a broad spectrum of news categories ranging from unscheduled news such as political events, natural disasters, war and conflict, etc., to scheduled news such as the release of important macroeconomic indicators or press conferences, as well as forecasts, discussions, and opinions of any of these news items.

For any news record that can be matched with an event category, RavenPack generates an Event Sentiment Score (ESS) signaling the potential impact of the news on the related entity (e.g., an economy or a financial instrument). For example, in terms of war and conflict, the entity is usually the country that is related. The same applies to economic growth, which refers to a growth in a certain country or region. These news are not specifically focused on crude oil. We also have crude oil specific news, for example, crude oil supply. The entity is then crude oil. A rise in supply means potential downward pressure for the oil prices. Therefore, such news are assigned a negative ESS. In [Table A.1](#), we provide some examples for the news categories examined in our paper.

The ESS analytic is determined by systematically matching stories typically categorized by financial experts as having short-term positive or negative financial or economic impact. The strength of the score is derived from a collection of surveys where financial experts rated entity-specific events as conveying positive or negative sentiment (direction) and to what degree (strength). Their ratings are encapsulated in an algorithm that generates a score ranging from 0–100.<sup>2</sup> In addition, RavenPack also assigns each news a novelty score and a relevance score. A novelty score of 100 (ranging from 1 to 100) documents the first story reporting a categorized event and is considered to be the most novel. The relevance score (ranging from one to 100) indicates how strongly related the entity is to the underlying news story. A value of 100 indicates that the entity plays a key role in the news story and the content is therefore most relevant for this entity.

We identify and work with event categories that are relevant for crude oil, both by type and the relevance and novelty scores. We choose those geopolitical and macroeconomic categories that are relevant for the supply and production of oil or that are related to the business cycle and can affect the consumption of crude oil. For each, we choose news events with both relevance score and novelty score of 100, representing the first mention of the news specific to a given event category. Furthermore, because our data set covers international news, for each category we are able to select those news based on its related entities, i.e. usually countries that

<sup>1</sup> More broadly, our paper is also related to the relationship between oil prices and the macro economy. See, for example, [Heath \(2016\)](#), [Hamilton and Wu \(2014\)](#) or [Bessembinder and Chan \(1992\)](#) for references.

<sup>2</sup> See the RavenPack News Analytics User Guide for more details. Here, it is worth mentioning that though we study the news impact on crude oil specifically, we include news on the other entities to make sure they cover a broad range of topics. More details can also be found in the first paragraph of Section 3.

**Table 1**

Summary statistics — crude oil returns and news sentiment categories. This table reports the summary statistics of daily crude oil returns and news sentiment scores of the listed categories. The sample period ranges from January 1, 2000 to March 31, 2013. Statistics of the news scores are reported by category. The penultimate row summarizes statistics of the news scores of the whole sample. The last row reports statistics of the daily crude oil returns. The news sentiment scores range from a minimum of −50 to a maximum of 50. Positive and negative scores stand for positive and negative tone of the news respectively. The score of 0 stands for neutral sentiment, which has been eliminated from our sample. Column 2 reports the number of news available for that category. Column 3–7 reports the mean, median, standard deviation, minimum and maximum of the news scores.

	Obs.	Mean	Median	Std. Dev.	Min	Max
Terrorism	120	−29.00	−25	5.68	−37	−25
War & conflict	367	−20.99	−15	9.79	−43	24
Civil unrest	2111	−14.25	−19	11.51	−47	27
Natural disasters	216	−9.63	−13	11.74	−21	25
Government	503	−15.73	−22	17.23	−35	29
Sovereign debt	132	−11.34	−32	29.12	−41	42
Public finance	1051	−6.54	−20	28.42	−50	50
Retail sales	1699	6.73	8	26.54	−50	50
Consumer confidence	834	1.16	20	32.73	−44	45
Foreign exchange	70030	0.70	33	32.99	−37	37
Housing	516	−0.92	−4	26.33	−43	43
Interest rates	386	−19.37	−39	33.93	−39	47
Treasury yield	2564	0.61	1	12.45	−50	45
Private credit	145	9.10	9	8.44	−24	24
Employment	4035	3.45	9	38.10	−50	50
Consumer spending	395	1.63	5	32.40	−50	50
Durable goods orders	373	1.63	7	33.84	−50	50
Recession	206	−43.00	−43	0.00	−43	−43
Economic growth	1217	31.48	45	32.04	−45	45
GDP growth	4383	18.43	45	40.12	−50	50
CPI	2333	8.50	37	34.58	−50	50
PPI	281	14.00	30	25.63	−30	30
Exports	5711	3.60	23	37.35	−50	50
Trade balance	4344	−1.97	−18	31.47	−50	50
Supply	750	−24.37	−22	10.66	−50	31
Demand	313	−15.85	−20	18.93	−33	33
Price target	249	4.98	24	23.60	−44	47
Drilling & pipeline accident	133	−1.92	18	23.99	−35	32
Total news	105397	1.29	9	33.26	−50	50
Crude oil	3313	0.0003	0.0012	0.0226	−0.1654	0.1334

**Table 2**

Category classification. This table reports the 28 news categories examined in the paper. The categories are reported mainly based on the RavenPack classifications. We group the news into three broader classes: geopolitical news, macro news, and oil supply and demand news.

<b>Geopolitical News:</b>				
Terrorism	War & conflict	Civil unrest	Natural disasters	Government
<b>Macro News:</b>				
Sovereign debt	Public finance	Retail sales	Consumer confidence	
Housing	Interest rates	Treasury yield	Durable goods orders	
Consumer spending	Recession	Economic growth	GDP growth	
CPI	PPI	Trade balance	Exports	
Foreign exchange	Employment	Private credit		
<b>Oil supply and demand:</b>				
Crude oil supply	Crude oil demand	Price target	Drilling & pipeline accident	

are the main driver of the global economy and hence the main oil consumers as well as the main oil producers. We elaborate on this point in Section 2.4. Our sample ranges from the inception of the data set is January 1, 2000 through March 31, 2013. We provide an overview of the data and the categories in Tables 1 and 2 and provide more discussions on them in Sections 3 and 4.

The news arrival is time-stamped to the second by RavenPack using a UTC (GMT) format. We convert this time stamp to EST or EDT time, which is the timing convention of crude oil futures data from the New York Mercantile Exchange (NYMEX). We also adjust for the switch between daylight saving and standard times each year.

## 2.2. Crude oil price data

We take WTI crude oil futures contracts traded on the NYMEX as our proxy for the crude oil market. We collect price, volume, and open interest data from the Commodity Research Bureau (CRB). In order to obtain a continuous time series, we take the nearest-to-maturity contract with maturity month  $T$  and roll to the next contract at the end of month  $T-2$  to avoid illiquidity issues near expiration. We then splice the log returns of the chosen contracts into a continuous time series.

According to the definition of the CME Group,<sup>3</sup> settlement prices of WTI crude oil futures are set as the volume-weighted average price of trades occurring on Globex between 14:28 and 14:30 ET. Being conscious of the dissemination time of the news, we conservatively assume when an event occurs before 14:20, investors can trade at the settlement price on the same day ( $t$ ). If a news report is released after 14:20, traders can only trade on the next day ( $t+1$ ) at that day's settlement price. News released over weekends or on holidays is treated as information available on the next working day before 14:20.

### 2.3. News classification

As part of our initial processing and analysis of the news analytics data, we identify three aspects that are relevant for signing the potential impact on oil prices. First, our data can be classified into three categories: event driven, economic data releases, and opinions/comments. Event-driven news is normally a timely report on an important event such as the breakout of a war or an incident. These news reports appear on the day that these events take place. Economic data releases are surveys or official releases of economic statistics. Similarly, these news occur on the day of the economic release. Surrounding either type of release, opinions and comments appear in the media, which may emphasize the importance of the information and/or contain additional information that could predict the possible outcome.

For example, war and conflicts may contain both event-driven news and news on further evolution and potential impact analysis. News about economic growth or recessions is often about opinions and comments about current economic conditions. Based on these types, the impact of the news on oil prices may also be different. Specifically, we find that for event-driven news, oil returns are significantly correlated with the contemporaneous news scores at the daily level, whereas news on economic growth, which contains opinions and comments, embody differential information than the statistical releases and can predict future oil prices. In fact, some news indices that have the corresponding statistical release available can predict the future outcome of the statistics. We elaborate on this issue in Section 4.9.

The second important aspect of the data is that for some categories there exists an important difference between the major oil-producing and oil-consuming countries. For example, the currencies of the major oil-producing countries typically move conversely to countries of major oil-consuming currencies. Coincidentally, the latter happen to be among the funding currency countries as well in the currency carry trade (see, for example, Ready (2013)). The same goes for exports or sovereign debt. Civil unrest in the major oil-producing countries, which is often associated with subsequent changes of regimes or governments, typically leads to temporary oil supply disruption. As a result, short-term oil prices rise. In contrast, civil unrest in major oil-consuming countries is mainly perceived as a sign for potential economic contraction and is correlated with an oil price decrease. On the one hand, major oil-producing countries earn higher revenue through high oil prices, consequently the sovereign debt level is expected to be reduced; on the other hand, rising oil prices are bad news for major oil-consuming countries, which raises their production and consumption costs and are usually accompanied by the coming increase of sovereign debt (Breunig and Chia, 2013). We detail on the difference between the major oil-producing and -consuming countries in Sections 2.4 and 4.4.

For geopolitical events, major oil-producing and oil-consuming countries play an important role, as their supply and demand make up a large part of global oil production and consumption, respectively, and hence exert a greater impact on oil prices. For economic events, industrial and export-oriented countries are more relevant due to their oil demand for production and consumption. Because their economic development is highly dependent on the world economy, these countries are quite sensitive to the global business cycle.

Finally, there is a clear difference between supply and demand side news. The impact of supply-side news in our sample is often severe but short-lived, e.g., concerns over Li by a and Iran. In contrast, we find that news from the demand side tends to have larger medium to long-term impact, e.g. rising demand from emerging markets and quantitative easing. These facts confirm the findings in Kilian (2009).

### 2.4. Classification of major oil-consuming and producing countries

As mentioned in the introduction, there are subtle differences between major oil-consuming and producing countries. Our procedure of classifying these two groups is as follows: First, from the U.S. Energy Information Administration (EIA) we obtain the annual total oil supply and total petroleum consumption per day by countries as proxies for the oil supply and consumption for the individual countries. The sample period is from 2000 to 2013. We then rank the ten largest consuming countries based on their total petroleum consumption and the fifteen most important oil-producing countries based on their total oil supply. There are countries who are in both of these two groups, meaning that they are the largest consumer and supplier of crude oil at the same time. For such cases, we calculate the consumption/supply ratio (CSR) by dividing the total oil consumption by the total oil supply of the related countries. To make sure that the consumption effect dominates the supply effect, we fix the cutting point of CSR at 1.2.<sup>4</sup> Those who has a CSR of larger than 1.2 is defined as oil-consuming countries and those lower than one are oil-producing countries. Our definition of major oil-consuming and producing countries can also be easily understood as oil-importing and exporting countries.

Though this imperfect classification may ignore the contributions of the related countries for the alternative group, however, for the discussed categories in this paper that are subject to the difference of these two groups such as civil unrest or sovereign debt, the effect of these countries either as producer or consumer dominates the opposite one. For example, U.S. and China are both

<sup>3</sup> <http://www.cmegroup.com/market-data/files/cme-group-settlement-procedures.pdf>.

<sup>4</sup> Our results are also robust with cutting point at 1.1 or 1.

**Table 3**

Major oil-consuming and -producing countries. This table reports the major oil-consuming and -producing countries defined in our paper. The major oil-consuming countries are the ten largest consuming countries based on their total petroleum consumption. The major oil-producing countries are the fifteen most important countries based on their total oil supply. The total oil supply and total petroleum consumption per day by countries are used as proxies for the oil supply and consumption and are obtained from EIA. For countries in both groups, we calculate the consumption/supply ratio (CSR). Countries with a CSR of larger than 1.2 is considered as major oil-consuming countries.

<b>Major oil-consuming countries:</b>				
U.S.	China	Japan	India	Germany
Eurozone	Korea	France	U.K.	Italy
<b>Major oil-producing countries:</b>				
Saudi Arabia	Russia	Iran	Mexico	Canada
United Arab Emirates	Venezuela	Norway	Kuwait	Nigeria
Iraq	Brazil	Algeria	Libya	Angola

the most important oil-producing and -consuming countries. However, their importance as oil-consuming players dominates that of oil-producing in terms of sovereign debt or export, because crude oil is not the major component their domestic income or export. In contrast, small countries that rely heavily on crude oil income such as Norway are more sensitive to oil price conditions. In the later analysis, we show that this classification matters, because news from the oil-producing and -consuming countries such as civil unrest actually reflect the supply and demand side shocks. In some cases, news from countries that belong to only one group has a much larger effect than countries that are in both groups. Table 3 reports the country classification.

### 3. News sentiment indices

After transformation, the RavenPack news scores range from  $-50$  (most negative sentiment) to  $50$  (most positive sentiment), with zero as a neutral score.<sup>5</sup> The sign and scale of the news scores are assigned by RavenPack based on its impact on the associated entity. Under certain circumstances, the sign is not necessarily identical with the sign of the news impact on crude oil. For instance, the breakout of the Iraq War in 2003 imposes a negative impact on the economy. However, due to the fact that Iraq is an important oil producer, oil price increased dramatically in anticipation of the coming war.

We summarize 28 news categories that are relevant for the oil market. For each category, we calculate an aggregate daily news index  $news_t$ . There may be several news reports on the same day in the same category. We therefore average any non-neutral scores across these news item daily.<sup>6</sup> In Section 4.1 below, we also discuss the alternative construction methods to show that our results remain robust to them. Also, as already mentioned, within some categories, we find opposing signs between oil-producing and oil-consuming countries. For example, civil unrest in the largest oil-producing countries are often negatively correlated with oil price changes, meaning that the negative sentiment in these countries is expected to reduce oil supply and hence lead to price increase. On the contrary, unrest in oil-consuming countries are typically associated with contraction in the economy. Hence, the sentiment scores in these countries are positively correlated with oil returns. In those categories in which oil-producing countries have systematically opposite signs than the major oil-consuming countries, we subtract the score of the oil-producing countries from those of the major oil-consuming countries before averaging.

Table 1 describes the data. It provides summary statistics on the 28 news indices as well crude oil returns. The standard deviation of the news scores ranges from zero (for recession indicators) to 40.12. Some categories such as terrorism and civil unrest exhibit predominantly negative scores. In total, 105397 news items are included in our sample. The sample mean is 1.29, which is close to neutral. The standard deviation of the whole sample is as large as 33.29. We standardize these scores to work with zero mean and unit standard deviation regressors in the empirical analysis below. The correlation among the news indices are relatively low. As Tables A.2 and A.3 show, the average correlation among the daily news indices is 0.002, with the maximum of 0.08 and minimum of  $-0.05$ . The average correlation among the monthly news indices is 0.02, with the maximum of 0.50 (between GDP growth and recession) and the minimum of  $-0.35$  (supply and employment). The parsimonious correlations among the indices eliminate our concern that these news indices may be driven by some common underlying state factors. To help readers understand the relationship between the important component of the news, especially among entity, news category, and the ESS score, we provide a sample of news in Table A.1.

## 4. Empirical results

### 4.1. Regression analysis and variance decomposition

The 28 news categories can be further classified into the following three groups: macroeconomic news, geopolitical news, and crude oil specific supply and demand related news. The details of the classification are described in Table 2. In this study, we are interested particularly in the different roles of macroeconomic and geopolitical news for crude oil.

<sup>5</sup> The original RavenPack scores range from 0 (most negative) to 100 (most positive), with 50 as the neutral score. We subtract 50 from the scores to make positive and negative news actually have the respective signs.

<sup>6</sup> We ignore neutral news as it is deemed uninformative one way or the other.



RavenPack assigns sentiment scores based on the potential impact of the news on the associated entity (e.g., an economy). In most cases, the news scores are therefore by construction positively related to the economic growth. Even so, for oil prices, the news coefficients may be negative, because oil prices react differently to different information about the economy. For example, during periods of war and conflict, terrorism, or natural disasters, news sentiment typically reflects uncertainty and potential supply reduction, which leads to a price increase despite the fact that the news is bad for the economy. Therefore, we expect the news from these three categories are negatively correlated with oil prices. We construct civil unrest as the news scores from the major oil-consuming countries minus the news scores from the major oil-producing countries. As explained in Section 2.3, civil unrest in oil-consuming countries typically signals economic contraction whereas in oil-producing countries it is a potential threat of supply disruption, we expect the news to be positive related to crude oil returns. Similarly, we expect positive sign for government news, as positive news from government in oil consuming countries is considered favorable for economic development, while negative news from government in oil producing countries will challenge the oil supply. News on oil supply increase receive a negative ESS. Therefore, intuitively, news on oil supply and demand should be positively associated with oil returns.

If the market is efficient, the information should be incorporated into prices immediately, especially when the news is publicly released and broadly disseminated. Tetlock et al. (2008) find that information embedded in news stories is quickly incorporated into stock market prices. Since we are dealing with public news, we expect that prices change immediately after the news release. To test this, we regress the log oil returns  $\Delta Oil_t$  on the contemporaneous news index in each category  $News_{i,t}$  using simple OLS. For robustness we add a set of control variables  $Control_{j,t}$  to the regressions: the past volatility of oil returns,<sup>7</sup> the oil futures basis at  $t-1$ , the log changes of trading volume and open interest at  $t-1$ , and monthly dummies  $M_t$ , respectively on the right hand-side of the regression equation:

$$\Delta Oil_t = \alpha + \beta \cdot News_{i,t} + \sum_j \varphi_j \cdot Control_{j,t} + M_t + \epsilon_t \quad (1)$$

Table 4 reports the results for daily data. As daily data contains substantial noise, especially for certain macroeconomic categories, we also examine the news impact at the monthly level in Table 5. The statistical significance is based on  $t$ -statistics calculated with auto-correlation adjusted Newey–West standard errors. The first column reports the univariate regression results without controls. Column 2 shows that the magnitude and significance of the news coefficients remain essentially the same after controlling for the seasonality, liquidity and market micro-structure effects.

Geopolitical news such as terrorism, war and conflict, civil unrest and natural disasters exerts an immediate, statistically and economically significant impact on the oil returns. For example, a one-standard deviation decrease of the news score for war and conflict is associated with a 13 basis points gain in the daily oil return on average, which is larger than the unconditional mean of daily crude oil returns (four basis points). The negative sign of the four geopolitical categories is consistent at the monthly level albeit with a smaller magnitude. The news coefficients of natural disasters and government are significant at the five percent level. The exception is terrorism, which tends to be a short-term effect that washes out over the course of a month.

Among the macroeconomic news, foreign exchange rates covary strongly with oil returns both at the daily and monthly level, as the oil price is denominated in US Dollar. The correlation is statistically significant at the one percent level. Daily oil returns also co-move heavily with other economic indicators such as sovereign debt, public finance as well as employment and private credit. News scores of treasury yield and interest rates are positively associated with oil returns at the daily level. A decrease in the news ESS on interest rates (corresponding to raising the nominal interest rates) by one standard deviation is expected to slow economic growth and decreases the daily oil return by 12 basis points. However, at the longer horizon rising treasury yield pushes up the monthly return by 2 percentage points in light of the higher expected inflation. This magnitude is well above the average monthly oil return of 0.8 percent.

Oil returns are sensitive to demand news. A positive demand shock raises oil returns by 11.21 basis points ( $p$ -value  $< 0.01$ ) at the daily level and 3.1 percentage points at the monthly level ( $p$ -value  $< 0.01$ ). However, oil prices seem to negatively associated with the supply news and hence positively related to the oil supply. A negative supply shock of one-standard deviation moves the oil returns up by 9 basis points ( $p$ -value  $< 0.05$ ) at the daily level and 1.36 percentage points at the monthly level ( $p$ -value  $< 0.1$ ). This shows that supply could be endogenously driven by oil demand and prices. When oil demand and prices are high, oil producers tend to raise oil supply. This evidence also supports the finding of Kilian (2009) that oil price appreciation is demand driven.

In light of the strong co-movement between oil returns and exchange rates returns, it would well be that the explanatory power of the macroeconomic and geopolitical news stems from the associated changes in exchange rates. Therefore, as a second robustness check, we translate the dollar price of crude oil into units of Special Drawing Rights (SDRs).<sup>8</sup> The fourth column of Tables 4 and 5 present the results with oil returns in denominated in SDR. As expected, the coefficient magnitude of foreign exchange decreases from 37.25 to 25.20. The other categories remain unchanged or even increase in magnitude and significance. Furthermore, we also regress oil returns on both the foreign exchange news index and one of the other remaining categories at the same time. The difference is marginal as well. In general, the dollar effect does not substantially change the news results.

Another concern is that our results could be heavily influenced by the financial crisis period. In particular, when plotting oil prices along with macroeconomic or geopolitical indexes, see Figs. 4 and 5, the comovement during the financial crisis between 2008 and 2009 is very pronounced. As a third robustness check, we therefore control for this effect by excluding the crisis period from our

<sup>7</sup> Specifically, we demean the log oil returns and square the obtained residuals, we then sum up the squared residuals over the past 60 days. The results are robust to using alternative volatility measures, such as 30-day or 360-day sum of squared residuals.

<sup>8</sup> [http://www.imf.org/external/np/fin/data/rms\\_sdrv.aspx](http://www.imf.org/external/np/fin/data/rms_sdrv.aspx).

**Table 4**

Daily oil returns and news. This table reports the news coefficients from the regression of daily log oil returns on the news scores in each category, as described in Eq. (1). The bottom panel reports the control variables used in running each of the regressions. Column 1 reports the news categories. Column 2–9 report the news coefficient in each regression on the respective news category and the control variables indicated below. Oil returns are expressed in basis points. News scores are standardized to have zero mean and unit variance. Newey–West standard errors are used to calculate the *t*-statistical significance. Numbers with \*: statistical significance at 10% level; \*: 5% level; \*\*: 1% level.

	1	2	3	4	5	6	7	8
Terrorism	−12.43**	−12.42**	−12.79**	−12.72**	−12.08**	−12.05**	−12.38**	−12.27**
War & conflict	−13.04**	−12.57**	−12.78**	−12.69**	−12.70**	−12.16**	−12.47**	−12.30**
Civil unrest	10.46*	9.86*	10.01*	9.43*	10.55**	9.96*	9.81*	9.21*
Natural disasters	−12.25**	−11.86**	−13.81**	−13.65**	−12.62**	−12.23**	−14.31**	−14.22**
Government	10.22*	10.36*	11.38**	11.69**	9.49*	9.69*	11.60**	11.96**
Sovereign debt	9.12**	9.34**	5.93*	6.00*	8.19*	8.45*	4.60	4.73
Public finance	13.51**	12.90**	10.58**	10.47**	13.69**	13.12**	10.88**	10.85**
Retail sales	13.97**	14.21**	11.79**	12.32**	14.50**	14.74**	12.04**	12.58**
Consumer confidence	4.15	4.56	1.98	2.28	3.45	3.82	1.27	1.53
Foreign exchange	−34.48**	−34.71**	−29.27**	−29.51**	−29.02**	−29.28**	−23.97**	−24.23**
Housing	−0.54	−0.18	2.51	3.06	−0.78	−0.43	2.23	2.81
Interest rates	11.91**	12.56**	6.67*	7.00*	11.70**	12.40**	6.05	6.43*
Treasury yield	10.99**	10.64**	9.21**	8.92**	10.38**	10.05**	8.43*	8.12*
Private credit	11.95**	12.14**	14.51**	14.73**	11.45**	11.61**	12.95**	13.18**
Employment	13.78**	13.66**	8.33*	9.10*	13.75**	13.66**	8.57*	9.33*
Consumer spending	2.42	1.79	0.78	0.76	1.34	0.70	0.08	0.02
Durable goods orders	−2.43	−2.74	−3.02	−3.27	−3.00	−3.29	−3.32	−3.55
Recession	2.71	1.20	−2.55	−2.57	1.63	0.17	−3.40	−3.37
Economic growth	−1.55	−2.56	−0.63	−0.50	−0.82	−1.84	−0.36	−0.31
GDP growth	5.45	4.05	1.01	1.29	5.05	3.66	0.93	1.20
CPI	3.22	2.96	4.90	5.23	4.18	3.92	5.42	5.74
PPI	−1.64	−2.39	−5.82	−4.58	−1.56	−2.24	−5.95	−4.53
Exports	4.06	3.43	6.42*	6.14	3.20	2.60	4.74	4.42
Trade balance	0.25	0.90	2.40	2.52	0.01	0.67	2.35	2.53
Supply	−9.09*	−9.23*	−5.58	−5.80	−9.89*	−9.98*	−5.81	−6.01
Demand	11.21**	10.64**	11.73**	11.99**	11.14**	10.59**	11.32**	11.57**
Price target	8.10**	7.08**	7.55*	7.36*	7.99*	7.00*	7.17*	6.93*
Drilling & pipeline accident	−13.30**	−13.65**	−13.34**	−13.45**	−13.34**	−13.73**	−13.36**	−13.52**
Control variables + monthly dummies		Yes		Yes		Yes		Yes
Without crisis period			Yes	Yes			Yes	Yes
SDR instead of Dollar price					Yes	Yes	Yes	Yes

sample. Specifically, we adopt the time-line of the financial crisis established by the Federal Reserve Bank of St. Louis.<sup>9</sup> To be more conservative, we exclude the period between February 2007 and August 2009 from our sample and run the same regressions on the rest of the sample. The third column of Table 4 shows that the coefficients of the geopolitical categories do not change much when we exclude the financial crisis period. Oil returns are even more sensitive to natural disasters. Government is the only category that exhibits a smaller though still highly significant coefficient. Given that government decisions were partially affected by the financial crisis, this result is not surprising. Among the macro categories in the third column of Table 5, the crisis period affects in particular recession and economic growth, and to a less extent CPI and PPI. Treasury yield and GDP growth are statistically significant at the five percent level. In addition, we also obtain similar results with the oil prices denominated in SDR or combining part or all of the three controlling methods at the same time. These regression results are shown in columns 5–8.

Furthermore, as the FX news make up a large proportion of our news sample, we carry out additional robustness checks to make sure that our results are not driven by the comovement between the other news categories and FX news. Table A.4 reports the results for both daily and monthly data. The results do not change much both in terms of the magnitude and statistical significance of coefficients. In fact, the correlation matrix in Tables A.2 and A.3 display low correlations between the FX news index and the other news indices.

In addition, in order to show that our news aggregation approach is robust to alternative methods and to avoid randomness in the ESS score assignment, we implement three additional robustness checks: First, following Dzieliński (2012) and Smales (2014), we adopt their method which also include the news with neutral ESS scores:

$$news_t = \frac{\sum_i \eta^p ESS_i^p + \sum_i \eta^z ESS_i^z + \sum_i \eta^n ESS_i^n}{\eta^p + \eta^z + \eta^n}, \quad (2)$$

where *p*, *z*, and *n* denote positive, neutral, and negative news.  $\eta$  is the number of the respective news. As a second method robustness check, we simply sum up the ESS scores of the same day instead of taking the average of it. Correspondingly, the monthly news score is the sum of all the news within that month. Third, to address the potential criticism that the ESS scores may be assigned by some randomness, we simply take the sign of the ESS and ignore its magnitude. With the dummies [−1, 0, 1] and following Eq. (2), we build the news indices.

<sup>9</sup> [http://www.imf.org/external/np/fin/data/rms\\_sdrv.aspx](http://www.imf.org/external/np/fin/data/rms_sdrv.aspx).



Table 5

Monthly oil returns and news. This table reports the news coefficients from the regression of monthly log oil returns on the news scores in each category, as described in Eq. (1). The bottom panel reports the control variables used in running the regression. Column 1 reports the news categories. Column 2–9 report the news coefficient in each regression on the respective news category and the control variables indicated below. Oil returns are expressed in percentage points. News scores are standardized to have zero mean and unit variance. Newey–West standard errors are used to calculate the *t*-statistical significance. Numbers with \*: statistical significance at 10% level; \*: 5% level; \*\*: 1% level.

	1	2	3	4	5	6	7	8
Terrorism	−0.09	−0.26	−0.09	−0.26	−0.09	−0.26	−0.09	−0.26
War & conflict	−0.30	0.01	−0.30	0.01	−0.32	0.01	−0.32	0.01
Civil unrest	−0.19	−0.89	−0.19	−0.89	−0.19	−0.89	−0.19	−0.89
Natural disasters	−1.79*	−1.85*	−1.79*	−1.85*	−1.79*	−1.85*	−1.79*	−1.85*
Government	2.12*	2.39*	2.12*	2.39*	1.88*	2.13*	1.88*	2.13*
Sovereign debt	−0.29	−0.14	−0.29	−0.14	−0.44	−0.28	−0.44	−0.28
Public finance	1.32*	1.06	1.32*	1.06	1.12	0.88	1.12	0.88
Retail sales	0.61	0.82	0.61	0.82	0.68	0.89	0.68	0.89
Consumer confidence	0.11	0.31	0.11	0.31	0.09	0.25	0.09	0.25
Foreign exchange	−2.68*	−2.88*	−2.68*	−2.88*	−2.18*	−2.35*	−2.18*	−2.35*
Housing	−0.46	−0.12	−0.46	−0.12	−0.47	−0.09	−0.47	−0.09
Interest rates	−0.23	0.12	−0.23	0.12	−0.21	0.16	−0.21	0.16
Treasury yield	2.15*	1.99*	2.15*	1.99*	2.02*	1.85*	2.02*	1.85*
Private credit	0.12	0.18	0.12	0.18	0.14	0.20	0.14	0.20
Employment	1.18	0.86	1.18	0.86	1.18	0.89	1.18	0.89
Consumer spending	1.60	1.19	1.60	1.19	1.45	1.01	1.45	1.01
Durable goods orders	−0.62	−0.90	−0.62	−0.90	−0.46	−0.70	−0.46	−0.70
Recession	2.95*	2.83*	2.95*	2.83*	2.69*	2.56*	2.69*	2.56*
Economic growth	2.02*	1.85*	2.02*	1.85*	1.91*	1.71*	1.91*	1.71*
GDP growth	2.02	2.35*	2.02	2.35*	1.96	2.33*	1.96	2.33*
CPI	1.66*	1.57*	1.66*	1.57*	1.57*	1.47*	1.57*	1.47*
PPI	−0.04	−0.48	−0.04	−0.48	−0.02	−0.32	−0.02	−0.32
Exports	1.10	0.76	1.10	0.76	1.12*	0.80	1.12*	0.80
Trade balance	0.43	1.04	0.43	1.04	0.45	1.09	0.45	1.09
Supply	−1.36*	−1.54*	−1.36*	−1.54*	−1.35*	−1.45*	−1.35*	−1.45*
Demand	3.10*	3.16*	3.10*	3.16*	3.11*	3.20*	3.11*	3.20*
Price target	4.43*	4.46*	4.43*	4.46*	4.15*	4.13*	4.15*	4.13*
Drilling & pipeline accident	0.80	0.63	0.80	0.63	0.76	0.55	0.76	0.55
Control variables + monthly dummies		Yes		Yes		Yes		Yes
Without crisis period			Yes	Yes			Yes	Yes
SDR instead of Dollar price					Yes	Yes	Yes	Yes

Table 6

Variance decomposition. This table reports the  $R^2$  explained by the regression of log oil returns on the contemporaneous news scores in the geopolitical (geo) and macro category respectively. Data of daily and monthly frequency are explored separately. In each case, the  $R^2$  and adjusted  $R^2$  are reported. The macro  $R^2$  is obtained by running a multivariate regression on all macro categories classified in Table 2. Similarly, the geo  $R^2$  is obtained with geopolitical categories. The total  $R^2$  is from the regression on both the macro and geopolitical categories. The common component share both by macro and geopolitical categories are distributed evenly to either of the two parts. The first row reports the value of  $R^2$  from the regressions above. The second row reports the proportion the  $R^2$  from macro or geo regression relative to the one of the regression with total  $R^2$ .

	$R^2$				Adj. $R^2$			
	geo	macro	geo + macro	Total	geo	macro	geo + macro	Total
Daily	0.0123	0.0367	0.0474	0.0550	0.0108	0.0312	0.0404	0.0469
pct	0.2425	0.7575	1		0.2485	0.7515	1	
Monthly	0.0856	0.2538	0.3113	0.4577	0.0556	0.1510	0.1870	0.3400
pct	0.2300	0.7700	1		0.2448	0.7552	1	

Results based on Eq. (1) are reported with both daily and monthly data. Table A.5 reports the results with news indices based on Eq. (2). Table A.6 reports the results with the sum of the news scores. Table A.7 reports the results with ESS dummies. Compared to Tables 4 and 5, the results and statistical significance are consistent. Therefore, our results are robust to the construction methods.

Having established that there are strong correlations between news and the oil price, we can turn to one of the main questions of the paper. Is macroeconomic or geopolitical news more relevant for explaining changes in the oil price? Or alternatively, how much of the total oil return variations can be attributed to each of these two news groups? To answer this question, we first run a regression with both the macro and geopolitical news indexes from the 24 categories (geo + macro) for the daily data. In Table 6, the adjusted  $R^2$  of this regression is 4.04 percent. We then regress oil returns on the geopolitical and macroeconomic indexes separately. The adjusted  $R^2$  are 1.08 and 3.12 percent, respectively, which correspond to 24.85 and 75.15 percent of the total return variations. We run the same type of analysis for monthly data. The proportion of the variance explained with the macroeconomic and geopolitical news is roughly the same (75.52% and 24.48%). Hence, macro economic variables tend to play a major in explaining the variation of oil returns both at the daily and monthly level, whereas geopolitical drivers, though the impact is large and immediate, play a second-order role.

Table 7

Monthly oil returns and news: Predictive regressions with control variables. This table reports the regression results from Eq. (3). Column 1 reports the news categories. Column 2–6 reports the coefficients  $\beta_k$  in the regression equation. Newey–West standard errors are used to calculate the  $t$ -statistical significance. Oil returns are expressed in percentage points. The  $p$ -values of the Wald-test whether the sum of the first two or all six lags are zero are reported in the last and third last column. News scores are standardized to have zero mean and unit variance. Numbers with \*: statistical significance at 10% level; \*: 5% level; \*\*: 1% level.

	$news_{t-1}$	$news_{t-2}$	$news_{t-3}$	$news_{t-4}$	$news_{t-5}$	$news_{t-6}$	$adj. R^2$	$\sum_{L=1}^6 \beta(L)$	Wald(6)	$\sum_{L=1}^2 \beta(L)$	Wald(2)
Terrorism	-0.39	-0.98	0.23	-1.22	1.95*	-0.78	0.10	-1.18	0.47	-1.36	0.62
War & conflict	-0.20	0.59	0.30	-1.08	0.57	1.81*	0.12	1.99	0.86	0.39	0.80
Civil unrest	1.07	0.95	-0.02	0.07	0.93	-1.03	0.09	1.96	0.98	2.02	0.35
Natural disasters	0.18	-0.74	1.58	-1.88*	0.69	-0.00	0.08	-0.16	0.82	-0.56	0.69
Government	1.82*	-1.70*	-0.88	0.07	-1.41*	0.44	0.15	-1.66	0.86	0.12	0.20
Sovereign debt	-2.28*	1.09	0.70	-1.26	-1.65	2.86*	0.21	-0.54	0.34	-1.19*	0.02
Public finance	-0.42	-0.93	1.77*	-1.71*	0.11	0.08	0.10	-1.11	0.57	-1.36	0.73
Retail sales	1.44*	-0.02	0.18	-1.69*	-1.98*	0.73	0.13	-1.33	0.47	1.42	0.35
Consumer confidence	-1.29	0.19	0.13	0.29	-0.36	1.52	0.08	0.49	0.99	-1.09	0.67
Foreign exchange	1.80*	-0.58	-0.55	-1.41*	-0.25	-0.26	0.10	-1.26	0.84	1.22	0.22
Housing	-0.57	0.48	-0.30	-0.12	-1.29*	1.05	0.08	-0.76	0.98	-0.09	0.82
Interest rates	0.13	0.37	-0.19	1.44	-0.27	0.81	0.08	2.28	0.93	0.50	0.93
Treasury yield	1.11	-0.55	-0.28	2.13*	0.57	-1.04	0.10	1.94	0.85	0.56	0.44
Private credit	0.96	0.04	-1.01	2.53*	-0.18	-1.15	0.13	1.21	0.33	1.01	0.68
Employment	0.68	-1.02	2.25*	-0.18	-0.00	-0.58	0.10	1.15	0.54	-0.34	0.77
Consumer spending	0.65	-1.11*	0.73	-1.62*	0.24	0.07	0.10	-1.05	0.97	-0.47	0.40
Durable goods orders	1.16*	-0.96	-0.64	0.28	-0.76	0.77	0.10	-0.16	0.99	0.20	0.45
Recession	2.55*	-0.38	0.59	-1.02	-0.90	1.19	0.12	2.04	0.24	2.17	0.13
Economic growth	1.48*	-0.33	0.72	1.61	-0.50	-0.20	0.11	2.78	0.76	1.15	0.36
GDP growth	2.88*	-0.09	-0.54	-2.55*	3.08*	-0.25	0.16	2.52	0.13	2.79*	0.10
CPI	2.61*	-1.86*	1.21	0.84	0.02	0.23	0.16	3.05	0.53	0.75*	0.01
PPI	0.69	3.18*	2.11*	-0.37	-2.86*	0.31	0.17	3.06	0.63	3.87	0.16
Exports	-0.66	0.33	-0.84	0.49	0.48	-0.11	0.05	-0.31	0.98	-0.33	0.77
Trade balance	-0.19	1.03	-0.83	1.48	-0.15	0.68	0.10	2.02	0.78	0.84	0.69
Supply	0.18	-0.31	-0.99	0.72	0.84	-0.93	0.06	-0.49	0.83	-0.13	0.97
Demand	-0.58	0.58	0.43	2.09*	1.05	0.28	0.14	3.85	0.73	0.01	0.65
Price target	-0.45	2.00*	-1.18	0.93	-1.17	0.14	0.08	0.27	0.78	1.55*	0.08
Drilling & pipeline accident	-0.44	0.82	-0.39	0.80	-0.70	1.69*	0.09	1.77	0.61	0.37	0.60

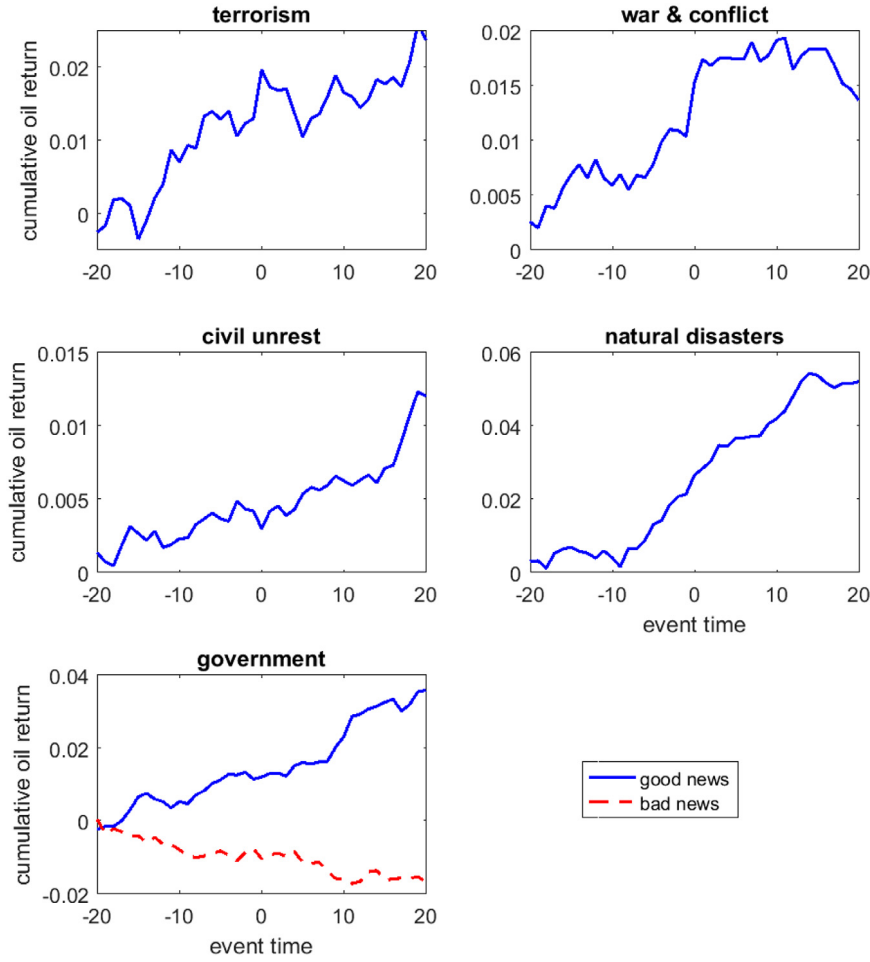
Though we document a strong correlation between news and oil prices both at the daily and monthly level, it is worth pointing out that we do not intend to impose a causal relationship between them. It could well be that both news and oil prices are endogenously driven by some common state variable. Instead, our goal is to attribute the oil return variations to different macroeconomic and geopolitical driven component and whether oil prices are driven by the state variables proxied by the specific news indices. Given that the correlations among the news indices are generally low, we could easily distinguish the different sources behind the oil return variation.

#### 4.2. Predictability

A number of papers on sentiment and stock returns have found that firms with positive news exhibit short-term momentum and subsequent mean reversion (Tetlock, 2007). Moreover, even old news, though not revealing any new information, appears to predict equity returns over the next one to five days (see for example Tetlock (2011)).

We employ event studies to illustrate the impact of the size and duration of these news on crude oil returns at the daily level, as the regression results based on daily data are not salient across the categories. We present the cumulative oil returns from 20 days before the news arrival to 20 days after. Fig. 1 presents the results for geopolitical news. The first four categories contain only or predominantly negative signs. News on terrorism and war and conflict leads to an immediate gain in oil price on the same day. Comparing the price at the end of the formation period to that at the beginning, the average oil return is about two percentage points. Civil unrest and natural disasters exhibit a more persistent impact on oil returns. Due to the longer duration of these events, oil returns rise already before the release of the respective news reports. The log oil prices increase by six percent on average at the end of the 20 days after the news event compared to the beginning of the testing sample. Positive government news leads to an oil return of 2.5 percentage points, whereas the impact of bad news is not obvious.

Compared to geopolitical news, macroeconomic news has a stronger and more smoothly evolving impact on oil prices. Fig. 2 shows the cumulative event returns for both positive and negative macroeconomic news sentiment. Oil prices rise gradually around positive GDP growth and economic growth news, and they decline around negative news. Consistent with Table 4, oil prices tumble in response to bad interest rate news (positive sentiment for the economy) and rise after good treasury yields news. For durable goods orders, cumulative oil returns are indifferent before the news release, but gain significantly in response to positive news compared to negative news. In terms of crude oil supply, the return difference at the end of the 20-day window is not very large. This result can be attributed to the fact that oil-producing countries tend to raise supply when the oil price increases dramatically, but the effect of the price decline is not immediately visible. In contrast, news about a supply decrease is associated with a monotone price increase. Intuitively, oil prices respond positively to increasing demand and the cumulative return 20 days after is around three percentage



**Fig. 1.** Event studies: Geopolitical impact. This figure illustrates the impact of geopolitical events on oil returns. We calculate cumulative log returns from 20 days before the news release to 20 days thereafter. The signs of the news index of terrorism, war & conflict, civil unrest, and natural disasters are negative, whereas government contains both positive and negative signs.

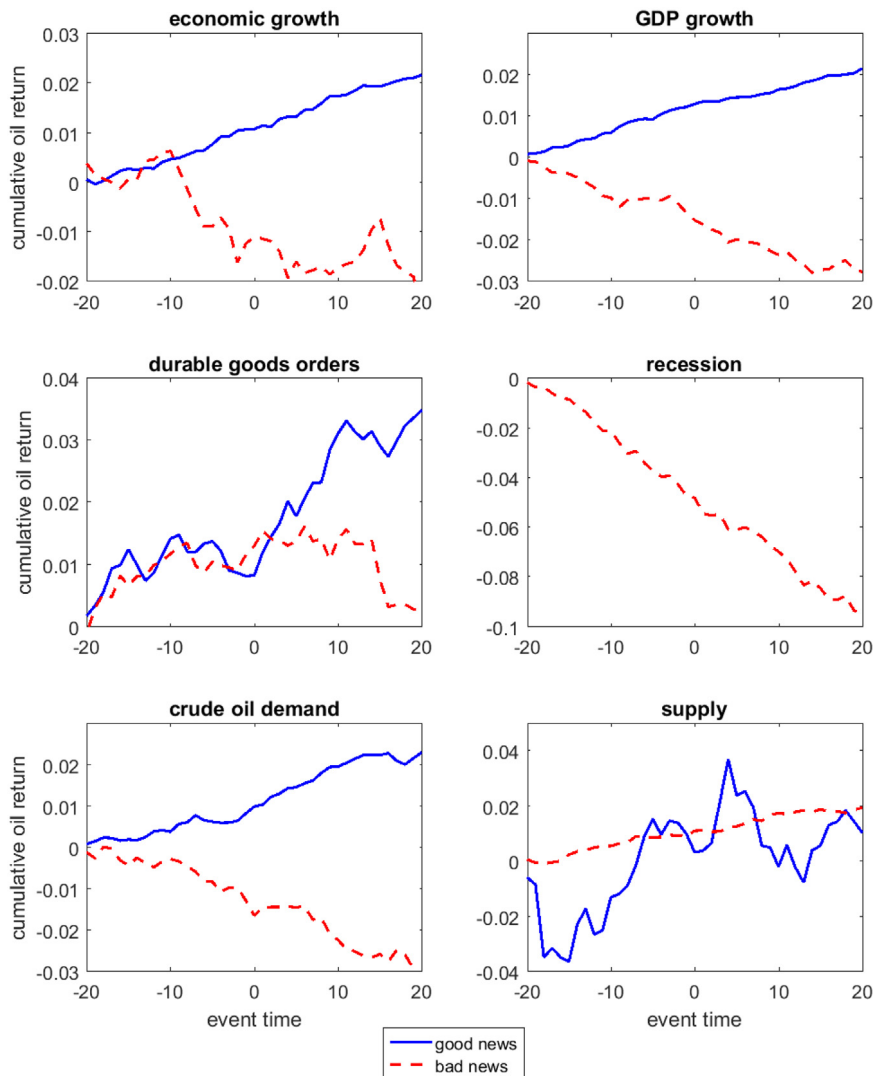
points compared to 20 days before. Price target news reflects the prognosis of large oil producers and financial institutions. Positive price target news mirrors the strong momentum effect of oil, while negative target news predicts further price declines.

We next test whether macroeconomic and geopolitical news have predictive power for the future oil returns at the monthly level. Specifically, we run predictive VAR regressions of oil returns on news, lags of news, lags of oil return, as well as the control variables mentioned above and the monthly dummies:

$$\Delta Oil_t = \alpha + \sum_{k=1}^K \beta_k L^k(News_{it}) + \sum_{k=1}^K \gamma_k L^k(\Delta Oil_t) + \sum_{k=1}^K \phi_k L^k(Control_t) + M_t + \epsilon_t. \quad (3)$$

$L^k$  is a lag operator, such that  $L^k(News_t) = News_{t-k}$ , where we let  $k=5$  for daily regressions and  $k=6$  for monthly regressions, potentially capturing a week or two quarters of persistence and delayed information processing, respectively.

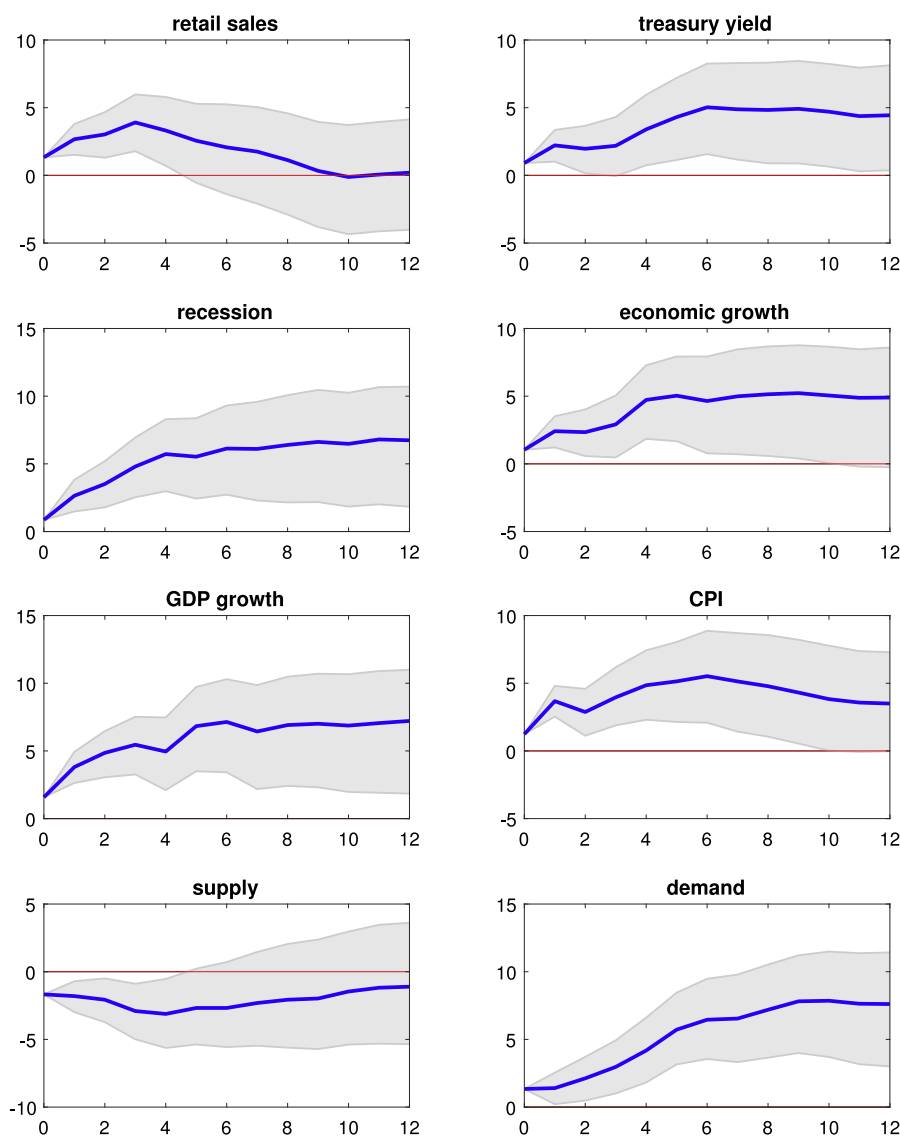
Table 7 shows that macro news, and in particular GDP growth, exhibits very strong predictive power for oil returns. An increase in the GDP growth news score by one standard deviation predicts an incremental log oil return of 2.90 percentage points ( $t$ -stat=2.3) over the subsequent month, which is both statistically significant at the five percent level and the scale of the coefficient is economically large. The Wald-test shows that the sum of the first two lags are statistically significant at one and 10 percent level. After excluding the financial crisis period in Table A.8, the predictive ability of GDP news remains robust. The coefficient of the first news lag increases both in magnitude and statistical significance (coef=2.9,  $t$ -stat=3.2). Other macroeconomic news categories such as recession, economic growth, CPI, and to a less extent, retail sales and durable goods orders also predict oil returns significantly at the first lag. It is worth mentioning that the PPI news index leads the CPI index and therefore predict oil returns two to three months in advance.



**Fig. 2.** Event studies: Economic impact. This figure illustrates the impact of economic events on oil returns. We calculate cumulative log returns from 20 days before the news release to 20 days thereafter. The solid line denotes news with positive sign. The dashed line denotes news with negative score.

The news variables also contribute positively to the adjusted  $R^2$  in most of the economic categories. As a benchmark, we first regress oil returns only on the lagged control variables without adding any news indices as in Eq. (3). This yields an adjusted  $R^2$  of 9.21%. Adding the individual news indices help increase the adjusted  $R^2$ . For instance, the adjusted  $R^2$  of sovereign debt is 24%, whereas that of the GDP growth is 17%. Thus, the explained proportion of variation with the additional news variable is much higher than the regression with only the control variables. From these results we conclude that oil returns are particularly driven by investor's bullish sentiment and positive expectation of future economic growth. Comparing the predictive power of geopolitical and macro news, we find that at monthly level, fundamental news about future economic growth dominates its geopolitical counterparts based both on the magnitude of the coefficients and the adjusted  $R^2$ .

Having shown the predictive results using VAR regression, it is interesting to examine further whether innovations in the news variables have an impact on the crude oil prices. The impulse responses are based on lower-triangular Cholesky decomposition of a VAR(6) fitted to the sequence news, oil returns, basis, volatility, open interest, and volume. To be consistent with the predictive regressions in Eq. (3), we deseasonalize the crude oil returns first before putting the deseasonalized time series with the other variables in the VAR equation. We present the cumulative impulse responses of oil returns to one standard deviation innovation in the news in Fig. 3. We also plot the 10% confidence band around the impulse responses. Consistent with the results from the predictive regression, one standard deviation shock in GDP growth leads to a cumulative increase of around 10 percent in the future 3 months, which is



**Fig. 3.** Impulse responses of oil returns to news shocks. This figure shows the cumulative impulse response of the deseasonalized oil returns to news of different categories. The impulse responses are based on lower-triangular Cholesky decomposition of a VAR(6) fitted to the sequence news, oil returns, basis, volatility, open interest, and volume. The gray shadow indicate 90% confidence intervals computed by block bootstrap. Data are monthly.

large and statistically significant. The other categories such as economic growth or exports demonstrates similar effect. An increase in supply has a negative impact on oil returns for the next 4 months, whereas demand impact is positive in the long run.

#### 4.3. Aggregate macroeconomic and geopolitical sentiment

In a similar fashion to [Cochrane and Piazzesi \(2005\)](#), we construct aggregate macroeconomic and geopolitical news sentiment indices as the weighted sums of the individual news indices within each group. Specifically, we weight each news category index by the coefficients of a multivariate regression of log oil returns on the group of macroeconomic or geopolitical news categories, respectively. [Figs. 4 and 5](#) respectively plot the geopolitical and macroeconomics sentiment indices along with the oil price. To smooth out some of the noise in the data we plot the standardized 120-day moving averages as opposed to the raw data series.

The geopolitical and macroeconomic news sentiment indices interact very differently with crude oil prices. The geopolitical index is characterized by very distinct political and natural crises. For example, the crude oil index co-moves strongly with the geopolitical index during the Syria, Egypt, Libya and Middle East unrest around the end of 2011 and beginning of 2012. Natural disasters such as hurricane or snow storm also contribute significantly to oil price increases. The macroeconomic index, in comparison, is smoother.





Fig. 4. Geopolitical news index and the crude oil prices. This figure coplots the aggregate geopolitical news index with the crude oil prices together. Both series are calculated as the accumulated returns over a moving window of 120 days, before being standardized to have zero mean and unit standard deviation.

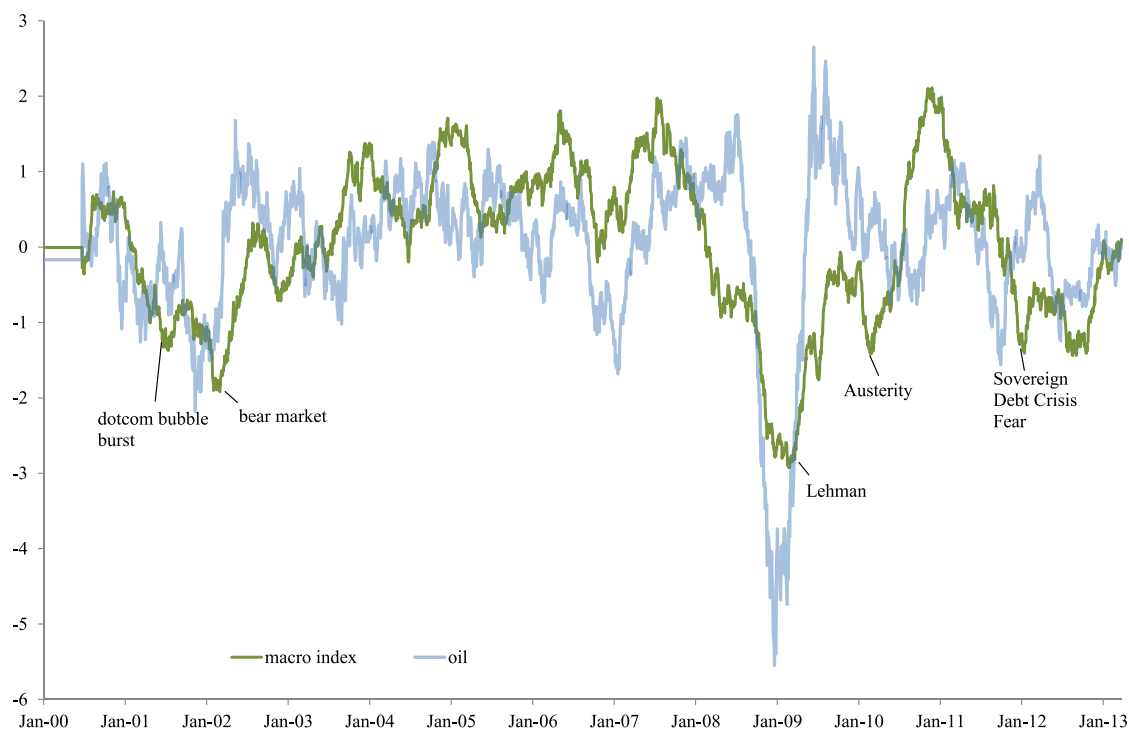


Fig. 5. Aggregate macro news index and the crude oil prices. This figure coplots the aggregate macro news index with the crude oil prices together. Both series are calculated as the accumulated returns over a moving window of 120 days, before being standardized to have zero mean and unit standard deviation.

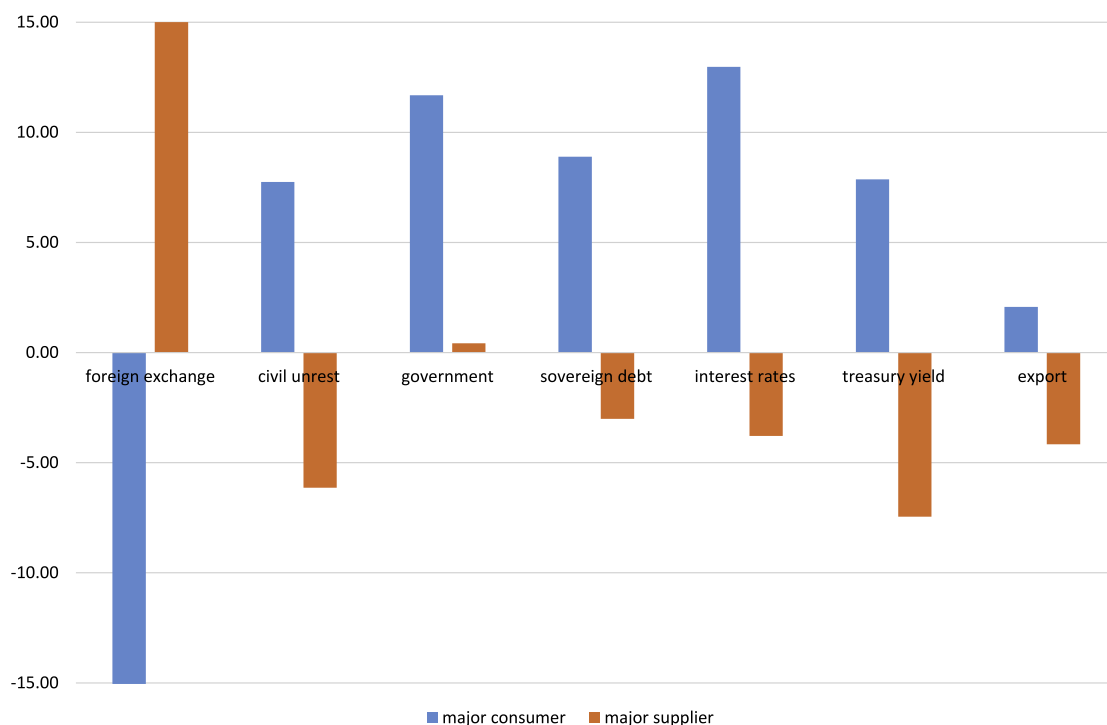


Fig. 6. Difference between the news impact from major oil-consuming and -producing countries. This figure plots the differential news impact from major oil-consuming and -producing countries on crude oil prices. The major oil-consuming and -producing countries are defined as in Table 3. Data are daily.

It reflects the global market economic growth during 2003 and 2007, the sub-prime crisis and the sovereign debt crisis. It is apparent that in mid of 2007 the macro economy enters into contraction, which leads the price collapse of crude oil in the second half of 2008. It is worth mentioning that we do not intend to impose a causal effect of news on crude oil prices. The purpose is to illustrate and distinguish the periods or variations where oil prices are rather driven by the macro or geopolitical underlying forces.

#### 4.4. News from major oil-consuming and -producing countries

As mentioned, we find subtle differences of news from major oil-consuming and -producing countries. To illustrate the difference, we plot the beta-coefficient in the daily regression analysis in Fig. 6. For some categories such as government or interest rates the difference between net major oil consuming and producing countries are even more accentuated.<sup>10</sup> The exchange rate of the major oil-producing countries comove positively with oil prices, whereas the contrary is found for the major oil-consuming countries. Civil unrest in consuming countries reflect the bad economic state and decrease in consumption, which affect oil prices negatively, while the same event in producing countries causes concern of oil supply cut which leads to positive effect. Expansive monetary policy such as decreasing interest rates is perceived as boosting oil consumption and results in rising oil prices in the major oil-consuming countries, whereas restrictive monetary policy in oil-producing countries will reduce investments and affect oil supply (rising oil prices). The difference between the beta coefficients of the two types of countries are, as reported in Table 4, is significant both in the economic and statistical sense.

#### 4.5. Trading volume and news sentiment

Traders in the market read and digest different quantities of news every day, which are then reflected in variations in trading volume over time. Theories predict contradictory signs of how the media impacts trading volume. We summarize three possible outcomes based on these theories. First, efficient market predicts no changes in trading activity. If we treat the content of the news as a public signal and market participants agree on how to interpret this signal and use it to update future expectations, disagreement is constant and the price of oil moves to a new equilibrium price without unusually high or low trading activity (e.g., Garcia (2013)). Still within a rational framework, Coval and Shumway (2001) suggest instead that changes in news sentiment are associated with higher

<sup>10</sup> For categories government and interest rate, news from the net major oil consuming and producing countries are more distinct from each other and have stronger predictive power. Therefore, in the analysis, we adopt this selection, whereas the results with the usual criteria are similar.

**Table 8**

Daily trading volume and news: Predictive regressions. This table reports the VAR estimates of daily log difference of oil trading volume on the lags of news sentiment in each category in Eq. (4). The coefficients  $\beta_k$ ,  $R^2$ , the sum of the coefficients of the five and two news lags  $\sum_{l=1}^l \beta(L)$ ,  $l = 2$  or  $5$ , and the  $p$ -value of the Wald-test whether the sum of the first two or all five lags are zero are reported in the last and third last column are reported. Newey–West standard errors are used to calculate the  $t$ -statistical significance.  $\Delta Vlm$  is expressed in percentage points. The data are daily. Numbers with  $\star$ : statistical significance at 10% level;  $*$ : 5% level;  $\ast$ : 1% level.

	$news_{t-1}$	$news_{t-2}$	$news_{t-3}$	$news_{t-4}$	$news_{t-5}$	$adj. R^2$	$\sum_{l=1}^5 \beta(L)$	Wald(5)	$\sum_{l=1}^2 \beta(L)$	Wald(2)
Terrorism	0.53	1.25	0.37	0.10	−0.50	0.17	1.76	0.98	1.78	0.42
War & conflict	−0.65	1.73 $\star$	0.52	−0.77	0.10	0.18	0.93	0.95	1.08	0.34
Civil unrest	2.40 $\ast$	2.35 $\ast$	0.32	−0.89	0.85	0.18	5.03	0.74	4.75 $\ast$	0.05
Natural disasters	0.69	1.25	−0.17	−1.44	−0.07	0.18	0.26	0.74	1.94	0.39
Government	−0.98	0.69	−1.37	−0.68	0.41	0.18	−1.94	0.94	−0.29	0.53
Sovereign debt	−1.65 $\star$	−2.40 $\ast$	−1.30	−1.35	0.50	0.18	−6.21	0.65	−4.05 $\ast$	0.04
Public finance	−1.05	0.30	0.53	0.93	0.48	0.17	1.19	0.99	−0.75	0.69
Retail sales	−0.62	−0.12	0.31	1.05	1.94 $\star$	0.18	2.56	0.94	−0.74	0.90
Consumer confidence	−2.41 $\ast$	−3.23 $\ast$	−4.10 $\ast$	−3.36 $\ast$	−1.42	0.18	−14.53 $\ast$	0.09	−5.64 $\ast$	0.01
Foreign exchange	0.88	0.02	0.85	−0.01	−1.40	0.17	0.34	0.98	0.90	0.85
Housing	−0.54	−1.76 $\star$	−1.96 $\ast$	−3.36 $\ast$	−7.40 $\ast$	0.21	−15.02 $\ast$	0.00	−2.30	0.20
Interest rates	−1.00	1.38	1.08	0.25	0.45	0.18	2.16	0.91	0.38	0.32
Treasury yield	1.63	0.16	−2.78 $\ast$	−0.84	2.87 $\ast$	0.18	1.04	0.38	1.79	0.45
Private credit	−10.74 $\ast$	−6.02 $\ast$	−2.46 $\ast$	−1.94 $\star$	2.63 $\ast$	0.24	−18.53 $\ast$	0.00	−16.76 $\ast$	0.00
Employment	0.07	−0.49	3.65 $\ast$	5.65 $\ast$	1.48	0.19	10.35	0.12	−0.42	0.95
Consumer spending	−7.06 $\ast$	−1.96 $\star$	−1.63	−0.56	0.53	0.20	−10.68 $\ast$	0.02	−9.02 $\ast$	0.00
Durable goods orders	−2.09 $\ast$	−2.46 $\ast$	−6.55 $\ast$	−4.82 $\ast$	−8.62 $\ast$	0.24	−24.54 $\ast$	0.00	−4.55 $\ast$	0.00
Recession	−0.50	−0.73	0.47	1.42	0.28	0.17	0.93	0.96	−1.24	0.82
Economic growth	−2.33 $\ast$	−2.47 $\ast$	−2.24 $\ast$	0.61	−1.51	0.18	−7.95	0.84	−4.81	0.20
GDP growth	−2.02 $\ast$	0.24	−0.07	0.19	1.86 $\star$	0.17	0.18	0.97	−1.79	0.61
CPI	−7.07 $\ast$	−4.29 $\ast$	−5.27 $\ast$	−3.53 $\ast$	−1.13	0.20	−21.29 $\ast$	0.00	−11.36 $\ast$	0.00
PPI	−1.61	0.18	4.89 $\ast$	2.87 $\ast$	3.47 $\ast$	0.19	9.80 $\ast$	0.02	−1.43	0.48
Exports	−0.39	−0.86	−0.16	0.92	−0.24	0.17	−0.73	1.00	−1.25	0.85
Trade balance	−0.24	−1.77 $\star$	−1.58	−0.15	−3.39 $\ast$	0.18	−7.14	0.74	−2.01	0.52
Supply	1.69 $\star$	−0.29	−1.94 $\star$	1.36	−1.52	0.18	−0.71	0.88	1.40	0.55
Demand	2.37 $\ast$	0.34	0.37	2.12 $\ast$	2.79 $\ast$	0.18	7.98	0.32	2.71	0.10
Price target	−0.53	−0.49	−0.53	0.11	0.22	0.17	−1.21	1.00	−1.01	0.79
Drilling & pipeline accident	−0.39	0.76	−1.26	−0.08	−0.15	0.17	−1.12	0.99	0.37	0.70

volatility and transaction costs, which in turn leads to a reduction in trading activity. Second, theories based on sentiment predict that noise traders and rational traders disagree when the tone of news is extremely positive or negative, as these types of news releases cause different interpretation and hence disagreement about the future price. As a result, there may be more trading (Antweiler and Frank, 2004). In line with this hypothesis, Tetlock (2007) discovers a nonlinear relationship between volume and media content, namely traders react to extremely pessimistic or optimistic news. As the sign of the news score is ignored, this proxy should measure the disagreement among investors. Third, if we can divide traders into informed and uninformed traders, then public information should eliminate information asymmetry between these two different groups and increase the willingness of uninformed traders to accommodate the liquidity shock induced by the informed traders from the last period (Tetlock, 2010). Therefore, volume should increase.

We formally test these types of hypotheses for the news categories in our sample by examining whether news release causes more trading volume in the following days. Following Tetlock (2007), we take the absolute value of the news score in our regressions. We estimate an autoregressive specification to examine the impact of news on trading volume over the following five days. We use the absolute value measure in our estimation:

$$\Delta Vlm_t = \alpha + \sum_{k=1}^K \beta_k L^k (|News_{i,t}|) + \sum_{k=1}^K \gamma_k L^k (\Delta Vlm_t) + \varepsilon_t. \quad (4)$$

Table 8 reports the results. In order to accentuate the difference between the two types of news, Fig. 7 illustrates the coefficient sum of the first two news lags. Geopolitical news such as terrorism, war and conflict, civil unrest, and natural disasters causes an increase in the oil trading volume over the next two to three days. The sum of the first to two or five lags are positive. One standard deviation increase in the absolute value of news leads to a total increase of 4.57 percentage points, which is statistically significant at the 1% level and economically large. In contrast, macroeconomic news such as sovereign debt, consumer confidence, durable goods orders, economic growth and CPI dampens trading volume after news release. For example, one standard deviation increase in the absolute ESS of CPI leads to a total decrease of 11.36 or 21.29 percentage points in the next two or five days. In general, the sum of the first two lags for almost all the macroeconomic news categories are negative.

The differential responses of oil trading volume to geopolitical and macro news suggest: (1) Geopolitical news seem to generate more uncertainty among investors and causes disagreement about the future oil price development, which leads to higher trading volume. (2) Macro news reduce trading afterwards. It appears that investors agree on how to interpret these news. The situation with the informed or uninformed traders as described in Tetlock (2010) is not particularly evident here. In this sense, the reaction of trading mirrors some kind of information efficiency in the macro area.

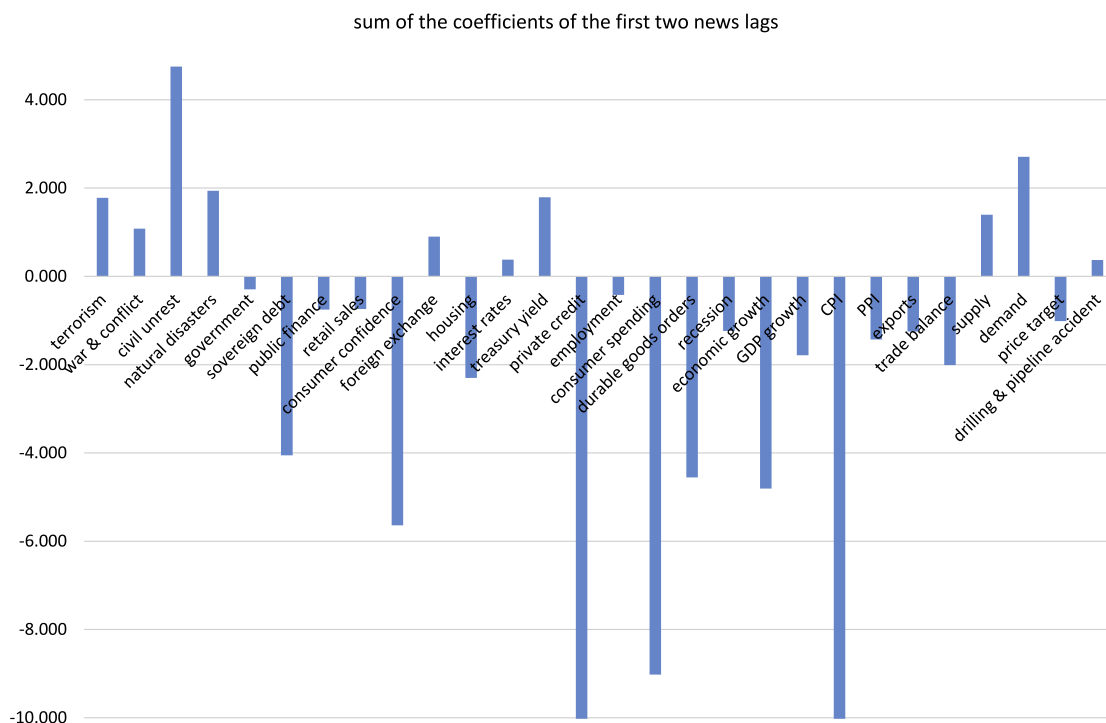


Fig. 7. News and trading volume of crude oil. This figure plots sum of the coefficient of the coefficients on  $|news_{t-1}|$  and  $|news_{t-2}|$  from the daily regression of Eq. (4).

#### 4.6. Relate news to oil demand, supply and inventory

Distinguishing the sources of oil demand and supply has been a much discussed topic in the literature on crude oil (see [Rapport \(2013\)](#) or [Kilian \(2009\)](#)). Based on the news characteristics, we might be able to classify them intuitively as demand- or supply-related news. For example, economic growth is most probably related to demand, whereas war and conflict may be associated with supply concern. However, the distinction is not absolute. For example, war and conflict could be a concern of supply disruption in the short run, but the economic damage and fear of recession thereafter are a demand effect. Therefore, without *a priori* judgment, we look at each news category separately and relate it to oil supply and demand. We obtain the monthly petroleum consumption in OECD countries as a proxy for oil demand, the world crude oil production as a proxy for oil supply, as well as the petroleum stocks in OECD countries as a proxy for oil inventory. All data series are from U.S. Energy Information Administration. In addition, we obtain the index of global real economic activity in industrial commodity markets (GEAI) from [Kilian \(2009\)](#). This index is used as a proxy for global demand in that paper. The correlation between the actual oil demand and supply is mildly positive (0.21).

We regress crude oil demand, supply or inventory growth, or GEAI at  $t$  respectively on each of the news indices  $news_{i,t}$ . In addition, for those categories with differential news impact from major oil-producing and -consuming countries we show the results separately. [Table 9](#) reports the beta loadings. Note that based on the ESS assignment, news on oil supply increase receive a negative ESS score, as it is supposed to have a negative impact on price. Therefore, the news index on supply is negatively correlated, and the coefficient is significantly different from zero at the 5% level. News on oil demand is also significantly related to the oil demand growth and negatively correlated with oil inventory growth.

The geopolitical news categories are not highly correlated with oil supply. Though the coefficient of terrorism is significant at the 10% level, it shows the “wrong” sign. The other categories such as war and conflict or natural disasters have a coefficient statistically indifferent from zero. However, war and conflict as well as natural disasters are negatively correlated with GEAI. One interpretation is that these political events, though the immediate impact is large, have little effect on the real oil supply in the long run. This result also partially explains or is in line with the monthly predictive results in [Table 7](#) that geopolitical news do not have predictive power for the long-run crude oil prices.

The macroeconomic categories are positive and statistically significantly correlated with GEAI. For instance, one standard deviation increase in news on GDP or recession is associated with a rise of 8–9 percentage points of GEAI, which is significant at the 1% level. News on treasury yield and export are closely related to oil production, implying that oil supply is endogenously related to monetary conditions and export activities. However, similar to the geopolitical news for production, the link between oil demand and macroeconomic news is weak.

Table 9

Compare news with crude oil fundamentals. This table reports the news coefficients from the regressions of crude oil consumption, production and inventory growth on the news indices. Column 6–9 report results with news from the major oil-consuming countries and column 10–13 with news from the major oil-producing countries for the categories, where news from both types of countries are employed. Newey–West standard errors are used to calculate the  $t$ -statistical significance. Consumption, production and inventory growth are expressed in percentage points. GEAI denotes the index of global real economic activity in industrial commodity markets from Kilian (2009). Data are monthly. Numbers with \*: statistical significance at 10% level; \*: 5% level; \*\*: 1% level.

	Consumption	Production	Inventory	GEAI	Consumption	Production	Inventory	GEAI	Consumption	Production	Inventory	GEAI
	All news				Major oil consumer news				Major oil producer news			
Terrorism	−0.15	−0.15*	0.01	−2.21								
War & conflict	0.13	0.04	−0.04	−3.71*								
Civil unrest	0.10	−0.03	−0.08	1.10	0.20	0.04	−0.09	0.04	0.07	0.08	0.01	−1.27
Natural disasters	−0.19	−0.06	0.13	−3.92*								
Government	−0.34	−0.08	0.05	1.52	−0.31*	−0.01	0.11	1.56	0.15	0.16*	0.09	−0.40
Sovereign debt	0.07	−0.04	−0.04	4.83*	0.18	−0.08	−0.04	3.54*	0.16	−0.05	0.02	−3.74
Public finance	0.12	0.02	0.04	−3.24								
Retail sales	−0.20	0.07	−0.06	−2.69								
Consumer confidence	0.11	0.01	−0.01	−0.62								
Foreign exchange	−0.01	−0.10*	−0.06	1.79	−0.05	−0.06	0.06	−0.38	−0.00	0.10*	0.10*	−2.30
Housing	−0.59*	−0.03	0.05	−2.88								
Interest rates	0.06	0.02	0.00	−2.50	0.04	0.03	0.02	−4.06	−0.06	0.07	0.06	−5.54*
Treasury yield	−0.14	0.14*	0.06	3.06	−0.08	0.05	0.07	−6.06*	0.10	−0.12*	−0.02	−6.79*
Private credit	0.04	0.07	−0.03	12.69*								
Employment	0.15	0.07	−0.05	−2.89								
Consumer spending	0.19	−0.06	−0.07	−0.70								
Durable goods orders	−0.23	−0.08	0.03	−4.49*								
Recession	−0.07	0.04	−0.02	8.38*								
Economic growth	−0.10	0.02	−0.05	4.22								
GDP growth	0.15	0.06	−0.08	9.28*								
CPI	0.14	−0.03	−0.15*	6.14*								
PPI	−0.07	0.05	−0.08	8.55*								
Exports	0.08	0.11*	−0.02	6.19*	0.02	0.08	−0.07	6.88*	−0.10	−0.04	−0.08	1.31
Trade balance	−0.01	0.03	−0.12*	4.33								
Supply	−0.05	−0.15*	−0.01	−2.91								
Demand	0.26*	0.09	−0.13*	3.54								
Price target	−0.34*	−0.05	−0.05	4.69*								
Drilling & pipel. accid.	0.04	0.03	−0.08	5.51*								

News from major oil producing and consuming countries are weakly related to oil supply and demand as well. News from major oil-consuming countries, which could be potentially regarded as a proxy for oil demand, are stronger correlated with GEAI than with demand or supply. In terms of news from the major oil-producing countries, only government or FX news are significantly linked to oil production, suggesting that government policy affect oil production and investment. The other categories do not demonstrate a strong link.

Alternatively, we also look at the aggregate news by taking the first components from news categories and compare them with oil supply or demand. The correlations between the news principle components and oil demand or supply is small as well. Furthermore, we run predictive regressions of oil production and consumption growth on the news data. The correlation of news with future oil supply or consumption growth is also weak.

The results above show that the direct news effect on crude oil demand or supply is not well prevalent, though news are closely related to the global demand and economic activities. One explanation may be that demand and supply are not elastic enough to respond to news. In this sense, the sentiment impact from news dominates.

#### 4.7. News impact on traders net positions

After looking at trading volume, we further examine how various market participants react to the news sentiment through variations in their net positions reported weekly to the U.S. Commodity Futures Trading Commission (CFTC). CFTC publishes Commitment of Traders (COT) reports which provide a breakdown of each Tuesday's open interest into two main groups: commercial and noncommercial traders. Commercial traders mainly use futures contracts for hedging purposes, while non-commercial traders are in the market to speculate. Since the net positions data is only reported weekly, we relate it to matching weekly observations of the sentiment indices and crude oil prices. Specifically, we match our news indices with every Tuesday's COT report. Here, similar to de Rooin et al. (2000) we define net long positions of a particular trader group  $i$  as:

$$NP_{i,t} = \frac{\#LongPositions_{i,t} - \#ShortPositions_{i,t}}{\#LongPositions_{i,t} + \#ShortPositions_{i,t}}. \quad (5)$$

Fig. 8 plots the net long positions of non-commercial traders and commercial traders calculated as in Eq. (5). The net positions of non-commercial traders exhibit stronger volatilities over time than those of commercial traders. Moreover, while non-commercial traders are mainly long ( $E(NP_t) = 0.12$ ,  $t$ -stat = 12.83) and commercial traders are on average short ( $E(NP_t) = -0.04$ ,  $t$ -stat = −15.56),



**Table 10**

Net noncommercial traders positions and news. This table reports the results from the VAR regressions of the net long noncommercial traders positions on the news indices as in Eq. (6). The coefficient  $\beta$ ,  $R^2$ , the sum of the coefficients of the five news lags  $\sum_{L=1}^5 \beta(L)$ , and the  $p$ -value of the Wald-test are reported. Newey–West standard errors are used to calculate the  $t$ -statistical significance. The net long noncommercial traders position is calculated according to Eq. (5) and is expressed in percentage points. The data are weekly. Numbers with \*: statistical significance at 10% level; \*: 5% level; \*\*: 1% level.

	$news_{t-1}$	$news_{t-2}$	$news_{t-3}$	$news_{t-4}$	$adj. R^2$	$\sum_{L=1}^4 \beta(L)$	Wald(4)	$\sum_{L=1}^2 \beta(L)$	Wald(2)
Terrorism	0.89*	0.44	0.01	−0.35	0.90	0.99	0.77	1.33	0.34
War & conflict	0.59*	0.08	0.13	−0.02	0.90	0.78	0.86	0.67	0.44
Civil unrest	0.00	−0.10	−0.09	−0.80**	0.90	−0.98	0.73	−0.09	0.97
Natural disasters	−0.13	−0.08	−0.08	0.43	0.90	0.13	0.79	−0.21	0.87
Government	−0.01	0.22	−0.08	0.37	0.90	0.50	0.96	0.21	0.85
Sovereign debt	−0.04	0.16	−0.05	−0.17	0.90	−0.10	0.99	0.12	0.91
Public finance	0.42	−0.13	−0.14	0.17	0.90	0.32	0.89	0.28	0.44
Retail sales	0.33	−0.01	0.35	−0.17	0.90	0.51	0.96	0.33	0.77
Consumer confidence	0.13	0.47*	0.83*	−0.84*	0.90	0.59	0.21	0.59	0.45
Foreign exchange	0.01	−0.22	0.05	0.45	0.90	0.29	0.88	−0.21	0.81
Housing	0.67*	−0.47*	0.52*	0.03	0.90	0.75	0.34	0.20*	0.04
Interest rates	0.20	−0.04	0.03	0.54*	0.90	0.73	0.68	0.16	0.78
Treasury yield	0.27	0.17	0.36	0.06	0.90	0.85	0.94	0.44	0.71
Private credit	−0.33	0.14	−0.58*	0.65*	0.90	−0.12	0.80	−0.19	0.80
Employment	0.52	0.41	0.31	0.29	0.90	1.53	0.85	0.93	0.46
Consumer spending	0.38	−0.09	0.20	−0.23	0.90	0.27	0.93	0.29	0.62
Durable goods orders	0.01	0.70*	0.29	−0.39*	0.90	0.62	0.72	0.71	0.28
Recession	0.41	0.27	0.27	0.19	0.90	1.13	0.83	0.67	0.44
Economic growth	0.49	0.54*	−0.02	0.23	0.90	1.24	0.74	1.04	0.29
GDP growth	−0.18	0.46	−0.19	0.73*	0.90	0.81	0.26	0.27	0.50
CPI	0.48	0.06	−0.32	−0.50	0.90	−0.28	0.90	0.54	0.67
PPI	−0.30	−0.12	−0.03	−0.03	0.90	−0.48	0.99	−0.43	0.71
Exports	0.21	0.04	−0.16	0.29	0.90	0.38	0.98	0.25	0.84
Trade balance	0.10	0.05	−0.29	0.09	0.90	−0.05	0.98	0.15	0.97
Supply	−0.30	−0.37*	−0.29	0.40	0.90	−0.56	0.70	−0.67	0.33
Demand	−0.07	0.66**	−0.10	−0.18	0.90	0.32	0.47	0.59*	0.05
Price target	0.44	−0.16	−0.25	−0.65**	0.90	−0.63	0.57	0.28	0.44
Drilling & pipeline accident	0.40	0.39	−0.06	−0.50*	0.90	0.23	0.38	0.78	0.23

both kinds of net positions vary around zero over most of the sample period, except for the recent period after January 2009. The correlation between the two time series is  $-0.89$ . For brevity, we examine the results of non-commercial traders in Table 10 and provide those of the commercial traders in Table A.9. The table reports results for the following regression specification:

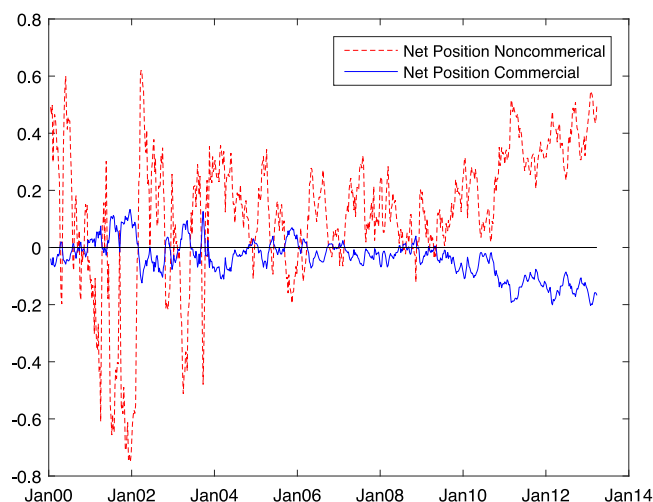
$$NP_{i,t} = \alpha + \sum_{k=1}^K \beta_k L^k (News_{i,t}) + \sum_{k=1}^K \gamma_k L^k (NP_{i,t}) + \sum_{k=1}^K \varphi_k L^k (Control_t) + M_t + \varepsilon_t, \quad (6)$$

where the control variables  $Control_t$  include as usual the past volatility of oil returns, the oil futures basis at  $t - 1$ , the log changes of trading volume and open interest at  $t - 1$ , and the monthly dummy variables. We examine the news impact of up to four lags (one month).

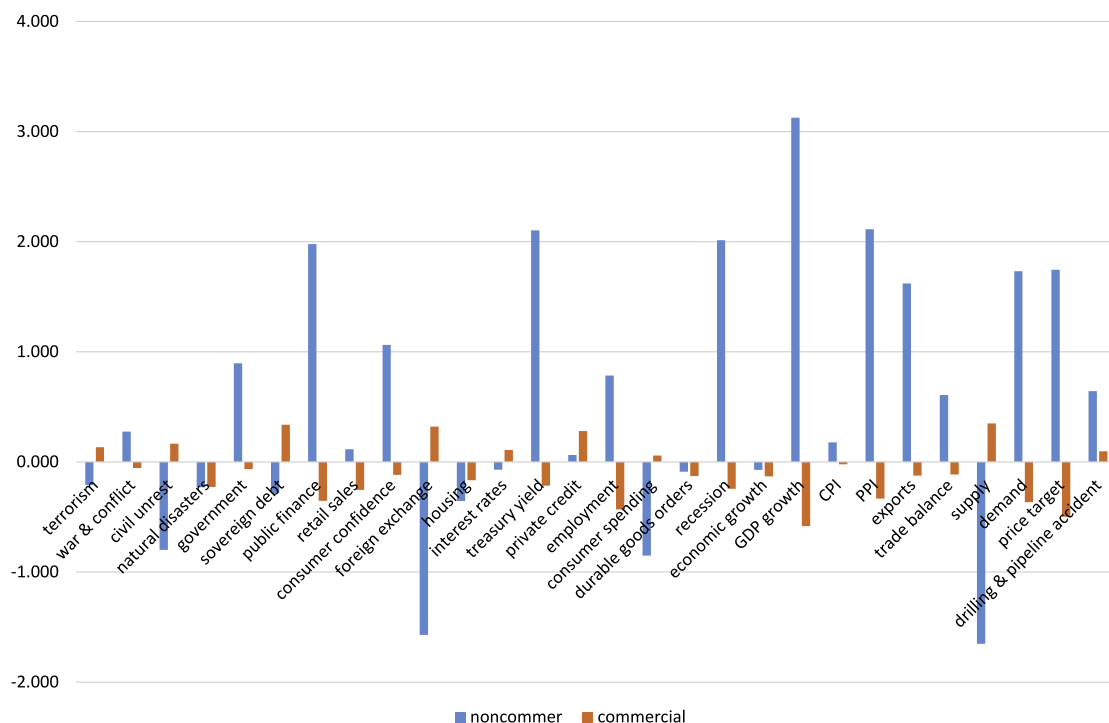
Following one standard deviation negative news sentiment related to terrorism, non-commercial traders reduce their net long positions in crude oil futures by 0.89 percentage point on average to avoid further uncertainty. A similar position reduction is observed for war or conflict news. In contrast, commercial traders go in the converse direction. Non-commercial traders increasing their long positions in face of growing demand for oil (the sum of the coefficients of the first two news lags are significant at the 5% level, whereas commercial traders on the other side of the contract increase their net short positions (decrease long positions). Non-commercial traders raise their long positions in response to most of the macro news. A one standard deviation increase of good news from the housing market leads to a growth in non-commercial net positions by 0.67 percentage points in the next week. The Wald-test shows that the sum of the first two news lag is significant at the five percent level for the last two categories. Fig. 9 further illustrates the differential response of traders regarding geopolitical news and macro news as well as the difference between noncommercial and commercial traders.

#### 4.8. News and volatility dynamics

As an alternative way to measure the effect of news sentiment on market uncertainty, we next relate the sentiment indices to the implied volatility  $IV_t$  from oil futures options, realized volatility  $RV_t$  calculated from tick-level data, and the variance premium  $VRP_t$ . The oil  $IV_t$  is a risk-neutral measure of expected volatility embedding the effects of both investor preferences and expectations. We obtain the oil options data from the Commodity Research Bureau. We construct the one-month implied volatility following Bakshi et al. (2003) from options on futures with two months to maturity. To calculate  $RV_t$ , we collect and clean high-frequency tick-level data for oil futures from TickData and then follow the procedure of Bollerslev et al. (2009). Specifically, daily realized variance is the sum of five-minute squared returns covering normal trading hours. Both realized and implied volatility are annualized. Finally,



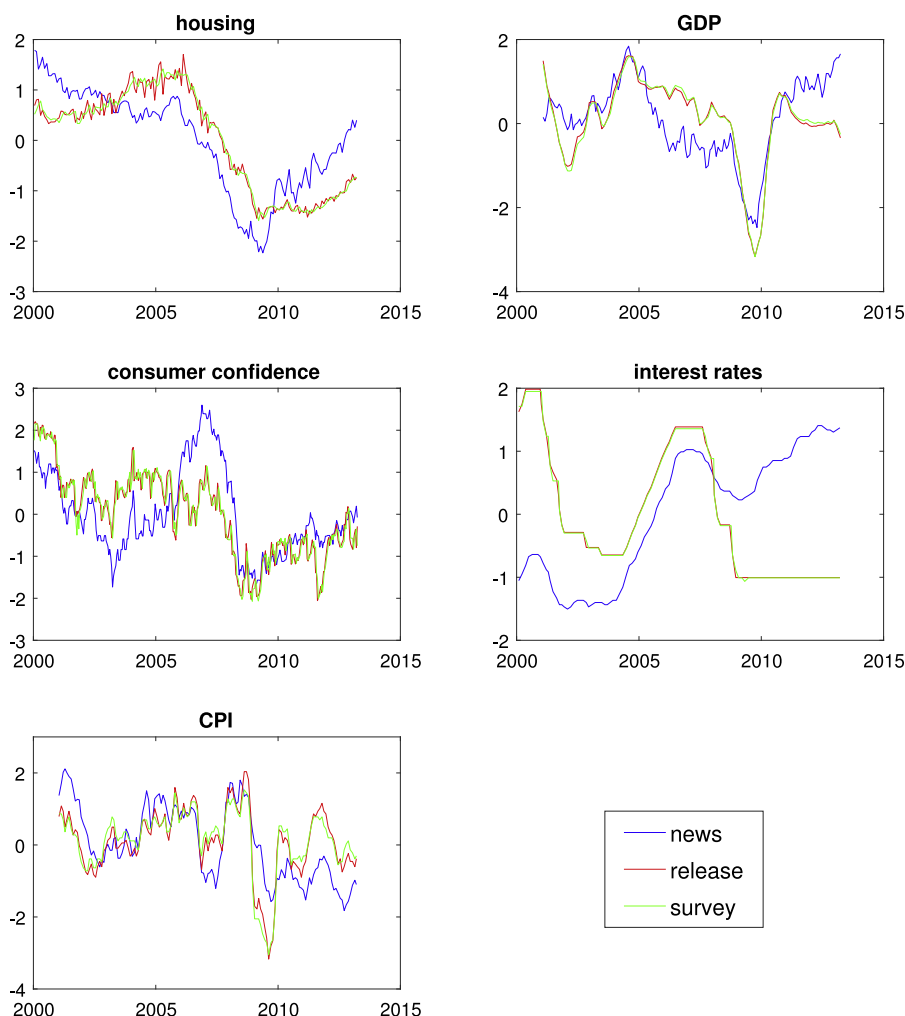
**Fig. 8.** Net positions of commercial and noncommercial traders. This figure coplots the net positions of commercial and noncommercial traders of the crude oil futures contracts calculated according to Eq. (5).



**Fig. 9.** Net noncommercial and commercial traders and news. This figure plots the news loadings from the contemporaneous regressions of the net long traders positions on the news variables. The regression equation is similar to Eq. (1), with the net traders positions on the left hand side. The net long traders positions are based on commercial and noncommercial traders of the crude oil futures contracts respectively. The news variables are standardized. Data are weekly to match the frequency of the COT data.

the volatility risk premium is defined as the square-root of the difference between implied variance and expected realized variance, where we simply use the average realized variance over the past month as proxy for expected realized variance.

We estimate the same specification as in Eq. (1) with the volatility variables in place of oil returns. In Table 11, the news coefficients of implied and realized volatility are negative for most of the categories, suggesting the counter-cyclical property of the oil volatilities.



**Fig. 10.** Data release and news sentiment. This figure illustrates the news indices (blue line) and the actual released statistics (red line). For comparison, the survey median from Bloomberg (in green) is also displayed. All series are standardized to have zero mean and unit standard deviation. . (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

In terms of the geopolitical news, when terrorism happens, one additional standard deviation of the negative sentiment score increases the annualized implied volatility by 0.26 percentage point, whereas the same increases the realized volatility by 0.39 percentage point. Remember that oil prices also rise regarding the uncertainty associated with negative geopolitical news. Positive macro news on the other hand dampens oil volatility.<sup>11</sup> Realized volatilities are sensitive to oil-related news such as supply and demand. News sentiment in both geopolitical (terrorism, civil unrest) and economic categories (public finance, retail sales, recession) is significantly associated with the oil implied volatility, suggesting that news imposes a significant impact on the expected variations in oil prices.

The volatility risk premium measures the additional cost investors are willing to pay in order to offset future uncertainty and is therefore often regarded as a risk aversion proxy. In our sample, the volatility risk premium is negatively correlated with the news indices, which means that the volatility risk premium is higher when the news sentiment is low. Therefore, investors are willing to pay more for the insurance in order to resolve the uncertainty. Good news reflect a decrease in investor's risk aversion, while bad news cause fear of future uncertainty.

<sup>11</sup> Foreign exchange news has a positive coefficient. The reason is that these news are typically negatively associated with the oil return (see Table 4).

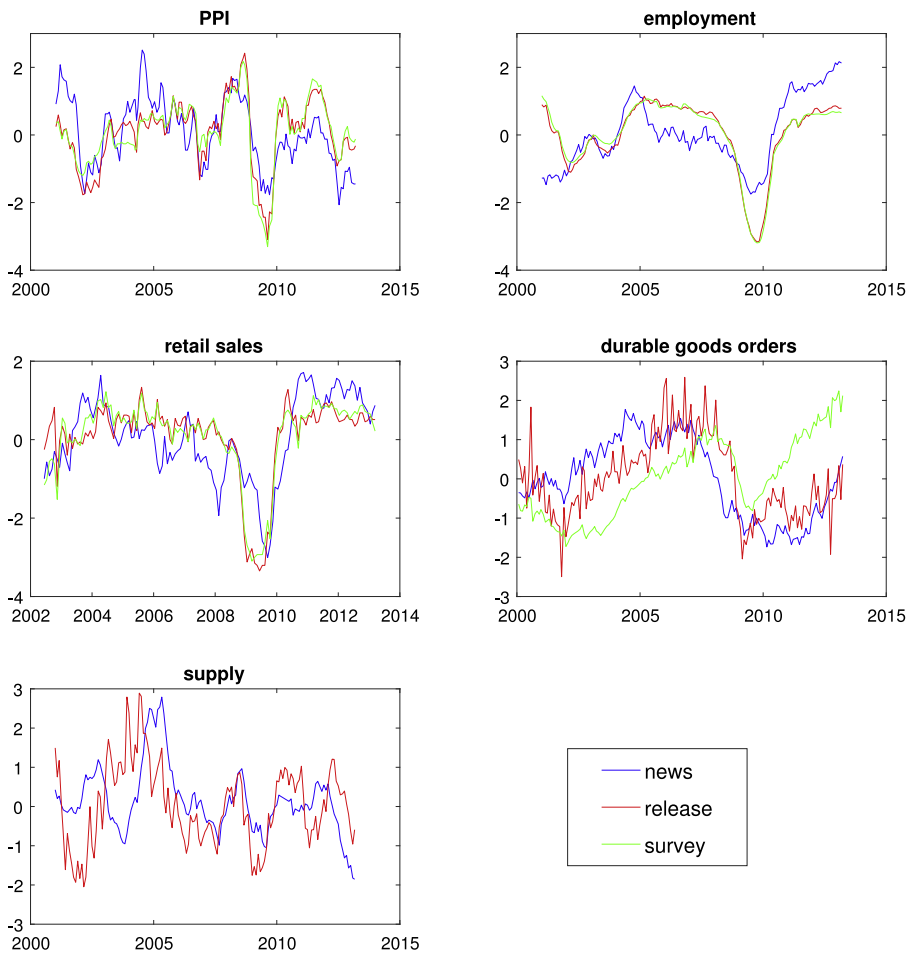


Fig. 10. (continued).

#### 4.9. Sentiment and scheduled news releases

Since news analytics contain not only “hard” news about the realization of macroeconomic statistics but also “soft” commentary about the market’s perception and predictions of the future trends in these series, it is natural to ask whether the sentiment indices predict future realizations of macroeconomic statics (beyond the information contained in previous releases).

We obtain the actual statistics and the time of release of those news categories that have such release available from Bloomberg. In addition, we employ the median of the survey on these indicators from Bloomberg for comparison with our news indices. The survey data in Bloomberg reflect to some extent the expectation of the financial experts and are widely used by investors thanks to their easy accessibility. We carefully match the news frequency with the frequency of the actual statistical release when plotting these series together or carrying out the regression analysis. To be conservative, we match the news up to day  $t - 1$  with the actual release in our analysis. For tractability reasons, we only examine US-related statistical release and news.

Fig. 10 plots the news sentiment indices, the actual released statistics, and survey for those categories with release data available. All time series have been standardized to ease comparison. In the figure, it is well recognizable that some news indices do lead the actual statistical release and the survey in some periods. For example, the housing index goes down around 2006 before the real statistics. The news index for durable goods orders peaks between early 2004 and early 2006, whereas the actual release data series peaks in January 2008. The news index starts to dip by the beginning of 2007, leading the actual data release which demonstrates the downward turn at the end of 2007. Gomes et al. (2009) show that stock returns of durable goods producing firms are much more exposed to cyclical changes and systematic risks than those of nondurable goods and service producing firms. This evidence is also reflected in the respective news sentiment. Similarly, the news index seems to lead the actual increase of interest rates (FOMC announcement). The news index goes down around 2001 and rises around 2004, earlier than the actual interest rate release. This is quite interesting, as we know that information on change of interest rates is kept confidential before the FOMC meetings (see, for example, Bernile et al. (2016)).

Table 11

Oil volatilities and news. This table reports the results from the regressions of oil return volatilities on the news sentiment in each category in a similar fashion as in Eq. (1). The examined oil volatilities are implied volatility from options on oil futures, realized volatility from high frequency data, as well as the variance risk premium, calculated as the difference of the first two variables. Newey–West standard errors are used to calculate the  $t$ -statistical significance. The data are daily. Numbers with \*: statistical significance at 10% level; \*: 5% level; \*\*: 1% level.

	IV			RV			VRP		
	$news_t$	$t$ -stat	$adj.R^2$	$news_t$	$t$ -stat	$adj.R^2$	$news_t$	$t$ -stat	$adj.R^2$
Terrorism	−0.26*	[−2.28]	0.76	−0.39*	[−1.88]	0.33	−0.13	[−1.24]	0.08
War & conflict	−0.13	[−1.23]	0.76	−0.37*	[−1.68]	0.33	−0.10	[−0.90]	0.08
Civil unrest	0.11	[0.96]	0.76	−0.07	[−0.32]	0.33	0.04	[0.34]	0.08
Natural disasters	0.06	[0.53]	0.76	0.22	[1.00]	0.33	−0.02	[−0.19]	0.08
Government	−0.03	[−0.28]	0.76	−0.34	[−1.53]	0.33	0.14*	[1.78]	0.08
Sovereign debt	0.00	[0.02]	0.76	0.02	[0.11]	0.33	0.07	[0.68]	0.08
Public finance	−0.22*	[−2.04]	0.76	−0.28	[−1.16]	0.33	−0.29*	[−2.91]	0.09
Retail sales	−0.23*	[−2.08]	0.76	−0.36	[−1.49]	0.33	−0.22*	[−2.27]	0.08
Consumer confidence	−0.12	[−1.13]	0.76	−0.05	[−0.23]	0.33	−0.09	[−0.94]	0.08
Foreign exchange	0.33*	[2.81]	0.76	0.57*	[2.20]	0.33	0.19*	[2.01]	0.08
Housing	0.05	[0.45]	0.76	−0.27	[−1.40]	0.33	0.13	[1.39]	0.08
Interest rates	0.06	[0.57]	0.76	0.10	[0.35]	0.33	0.07	[0.74]	0.08
Treasury yield	−0.04	[−0.38]	0.76	−0.09	[−0.38]	0.33	−0.04	[−0.54]	0.08
Private credit	−0.04	[−0.49]	0.76	0.05	[0.31]	0.33	−0.04	[−0.51]	0.08
Employment	−0.03	[−0.27]	0.76	−0.73*	[−3.18]	0.33	0.02	[0.22]	0.08
Consumer spending	0.09	[0.84]	0.76	−0.24	[−0.79]	0.33	0.01	[0.11]	0.08
Durable goods orders	0.09	[0.85]	0.76	−0.08	[−0.43]	0.33	−0.04	[−0.47]	0.08
Recession	−0.28*	[−2.08]	0.76	−0.48	[−1.46]	0.33	−0.12	[−1.29]	0.08
Economic growth	−0.05	[−0.47]	0.76	0.22	[0.91]	0.33	−0.05	[−0.65]	0.08
GDP growth	−0.09	[−0.76]	0.76	0.05	[0.18]	0.33	−0.07	[−0.61]	0.08
CPI	−0.05	[−0.49]	0.76	0.06	[0.26]	0.33	0.01	[0.09]	0.08
PPI	0.00	[0.02]	0.76	−0.12	[−0.46]	0.33	0.00	[0.04]	0.08
Exports	−0.04	[−0.36]	0.76	0.23	[1.00]	0.33	−0.17*	[−1.79]	0.08
Trade balance	−0.03	[−0.34]	0.76	−0.55*	[−2.17]	0.33	0.01	[0.09]	0.08
Supply	0.05	[0.53]	0.76	0.45*	[2.12]	0.33	−0.04	[−0.47]	0.08
Demand	−0.24*	[−2.09]	0.76	−0.45*	[−2.27]	0.33	−0.01	[−0.09]	0.08
Price target	−0.07	[−0.55]	0.76	−0.78*	[−2.71]	0.33	0.30*	[3.42]	0.09
Drilling & pipeline accident	−0.05	[−0.57]	0.76	0.28	[1.38]	0.33	−0.00	[−0.01]	0.08

Table 12

News and macroeconomic statistical release. This table reports the results from the regression  $actual_{t+1} = c + b \cdot variable_t + d \cdot actual_t + \varepsilon_{t+1}$ , where  $variable$  refers to  $news$ ,  $news_t$ , and  $survey_t$ , or  $survey_t$ . Newey–West standard errors are used to calculate the  $t$ -statistical significance. The data frequency is matched to the frequency of the statistical release. Numbers with \*: statistical significance at 10% level; \*: 5% level; \*\*: 1% level.

	$news_t$	$t$ -stat	$survey_t$	$t$ -stat	$adj.R^2$
Housing	0.06*	[3.87]			0.967
	0.06*	[2.71]			0.976
			0.65*	[8.42]	0.975
GDP	−0.04	[−1.07]			0.713
	−0.04	[−1.09]			0.711
			0.02	[0.09]	0.712
Consumer confidence	−0.01	[−0.55]			0.938
	−0.01	[−0.58]			0.938
			0.02	[0.50]	0.938
CPI	−0.00	[−0.42]			0.998
	−0.00	[−0.15]			0.999
			0.46*	[3.32]	0.999
PPI	0.02	[1.24]			0.995
	−0.01	[−0.66]			0.995
			0.23*	[2.46]	0.995
Employment	0.04	[0.99]			0.616
	0.06	[1.47]			0.643
			0.41*	[2.67]	0.642
Retail sales	0.07*	[2.18]			0.989
	0.05*	[1.69]			0.989
			0.33*	[2.39]	0.989
Durable goods orders	0.24*	[4.29]			0.668
	0.34*	[5.23]			0.684
			0.37*	[2.81]	0.635
Supply			0.03	[0.86]	0.958
	0.06*	[2.39]			

However, these news indices seem to be slow in picking up the upward trend in the aftermath of the drastic drop-down. In the case of durable goods orders, the news index is lags the actual data. It moves sideways between 2009 and 2011, while the actual data pick



up the increasing trend rapidly. Similar cases are found in other categories such as GDP, retail sales, or employment. This evidence suggests anecdotally that the news sentiment index may not be good at picking up the early economic recovery, as the pessimistic sentiment pervades for a longer period. However, it could be a good candidate to detect the turning point from boom to bust and is a useful indicator to forecast crisis.

The other news indices such as crude oil supply or employment track the behavior of the released data well although it is more volatile than the actual data. Given the fact that our news indices are available at the daily frequency, whereas the macro statistical release appears on the monthly or quarterly basis, it is natural to think that the news sentiment index contains more information than the released economic statistics. This is not surprising as it also incorporates expectation and evaluation of the data release and can reflect the respective economic situation more timely.

We provide more formal evidence by regressing the actual released statistics  $actual_{i,t+1}$  on its own lag  $actual_{i,t}$  and the first lag of the respective news index  $news_{i,t}$ . For comparison, we also add the survey data if available  $survey_{i,t}$ . Table 12 reports the results. Consistent with Fig. 10, housing, durable goods orders, retail sales, and supply positively predict the next period statistics. Even after controlling for the survey data, these results are still robust. This evidence suggests that news indices indeed contain some predictive power for the official data releases. The other news categories do not show predictability for the actual release, however, it is evident from Fig. 10 that they track the statistics very closely.

## 5. Conclusion

News about macroeconomic fundamentals and geopolitical events affect crude oil markets differently. Using sentiment scores for a broad set of global news of both geopolitical and macroeconomic types, we find that news related to macro fundamentals have an impact on the oil price in the short run and significantly predict oil returns in the long run. Geopolitical news have a much stronger immediate impact but exhibit not predictability. Moreover, geopolitical news generate more uncertainty and greater trading volume, consistent with a disagreement explanation, while macroeconomic news are associated with subsequent lower trading volume. News from major oil-consuming countries and producing countries can be conversely related to oil returns due to the dominating demand and supply effect they represent respectively. Finally, we find that news sentiment is useful in detecting the downward turning point and can partially predict the future realizations of economic data thanks to their information content about future expectations.

## Appendix

See Tables A.1–A.9

**Table A.1**

A news sample. This table provides a sample of news, which comprises the entity, category and ESS score, in order to help readers understand the relationship among the three items.

Category name	Entity	Category description	ESS
Terrorism	New York, NY, US	Terrorist-threat-target	–25
War & conflict	Iraq (IQ)	War-declaration-target	–43
Civil unrest	Libya (LY)	Civil-unrest	–50
Natural disasters	Iwaki, FUKUSHIMA-KEN, JP	Tsunami	–25
Government	Republic of Korea	Cabinet-resignation-authority	–22
Sovereign debt	Brazil (BR)	Sovereign-debt-down	42
Public finance	Germany (DE)	Government-budget-deficit	–50
Retail sales	United States (US)	Retail-sales-down	–50
Consumer confidence	United States (US)	Consumer-confidence-down	–50
Foreign exchange	U.S. Dollar	Currency-rate-depreciate	–33
Housing	United States (US)	Home-sales-new-down	–43
Interest rates	European Central Bank	Interest-rate-down	13
Treasury yield	Canada (CA)	Treasury-bill-yield-down	47
Private credit	United States (US)	Private-credit-down	23
Employment	United States (US)	Employment-down	–46
Consumer spending	Eurozone	Consumer-spending-down	–50
Durable goods orders	United States (US)	Durable-goods-down	–41
Recession	Eurozone	Recession	–43
Economic growth	United States (US)	Economic-growth-down	–45
GDP growth	United States (US)	GDP-down	–50
CPI	Germany (DE)	CPI-down	22
PPI	Eurozone	PPI-down	–50
Exports	Japan (JP)	Exports-down	–50
Trade balance	Canada (CA)	Trade-balance-deficit	–50
Supply	Crude oil	Supply-increase-commodity	–50
Demand	Crude oil	Demand-increase-commodity	44
Price target	Crude oil	Price-target-downgrade	–47
Drilling & pipeline accident	Saudi Arabia (SA)	Drilling	18

Table A.2

Correlations among the News Indices — Daily. This table reports the correlations among the news indices. Data are daily.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)	(22)	(23)	(24)	(25)	(26)	(27)	(28)
(1) terrorism	1.00																											
(2) war & conflict	0.04	1.00																										
(3) civil unrest	-0.00	0.02	1.00																									
(4) natural disasters	0.05	0.04	-0.02	1.00																								
(5) government	0.02	-0.01	0.03	-0.01	1.00																							
(6) sovereign debt	-0.01	0.02	0.01	-0.01	-0.03	1.00																						
(7) public finance	0.00	-0.03	-0.01	-0.03	0.00	0.03	1.00																					
(8) retail sales	0.01	-0.02	0.01	-0.03	0.03	0.03	0.00	1.00																				
(9) consumer confidence	0.02	0.01	-0.01	0.01	-0.00	0.02	-0.00	0.02	1.00																			
(10) foreign exchange	0.01	-0.01	-0.01	-0.02	-0.04	-0.02	-0.03	0.01	-0.03	1.00																		
(11) housing	-0.01	-0.01	-0.02	0.01	-0.03	-0.03	0.00	-0.01	0.01	0.00	1.00																	
(12) interest rates	-0.01	-0.02	0.01	0.01	0.01	-0.02	-0.04	-0.02	-0.03	-0.01	0.03	1.00																
(13) treasury yield	0.00	-0.01	0.04	0.01	0.02	0.02	0.02	-0.05	-0.01	-0.04	0.01	-0.01	1.00															
(14) private credit	0.00	-0.03	0.01	-0.00	-0.03	0.01	-0.01	0.03	-0.01	0.04	-0.01	-0.02	0.01	1.00														
(15) employment	-0.02	-0.02	0.02	-0.00	0.01	0.01	0.04	0.01	0.00	-0.01	-0.00	0.03	-0.00	0.03	1.00													
(16) consumer spending	-0.01	-0.01	-0.01	-0.02	-0.02	-0.02	-0.02	-0.00	-0.00	-0.01	0.00	-0.01	0.01	-0.02	-0.03	1.00												
(17) durable goods orders	-0.00	0.01	-0.00	0.01	-0.02	-0.00	0.00	-0.00	0.01	-0.01	0.02	-0.01	0.02	-0.02	0.01	-0.00	1.00											
(18) recession	-0.02	0.04	0.02	-0.01	0.00	0.00	0.01	0.02	0.02	-0.02	0.00	-0.04	0.01	-0.01	0.01	0.00	0.02	1.00										
(19) economic growth	0.02	-0.01	0.01	-0.01	-0.02	-0.02	0.01	0.01	-0.00	0.02	0.00	-0.00	0.01	0.02	-0.00	0.03	-0.00	-0.01	1.00									
(20) GDP growth	-0.00	-0.03	-0.02	-0.02	-0.01	0.03	-0.00	0.01	0.01	-0.03	0.00	-0.03	0.03	0.06	0.01	-0.01	0.03	0.08	0.04	1.00								
(21) CPI	0.02	-0.01	-0.00	0.00	0.01	0.01	0.01	0.02	0.00	0.00	0.02	0.00	0.02	0.06	0.01	0.01	-0.01	0.01	0.01	0.02	1.00							
(22) PPI	-0.01	0.01	0.02	0.01	0.01	-0.00	-0.01	-0.01	0.00	-0.00	-0.00	-0.01	0.02	0.00	-0.00	0.02	-0.00	0.00	0.00	0.00	0.04	1.00						
(23) exports	-0.01	-0.04	0.03	-0.02	0.01	0.04	0.00	0.00	-0.00	-0.04	0.01	0.02	-0.00	0.04	-0.01	0.01	0.01	0.03	0.03	0.05	0.02	-0.02	1.00					
(24) trade balance	0.01	-0.01	-0.02	0.01	0.00	0.03	-0.03	-0.00	-0.01	0.04	0.01	-0.00	-0.02	-0.03	-0.02	0.02	-0.03	0.03	-0.00	-0.02	-0.01	0.01	0.01	1.00				
(25) supply	0.02	0.01	-0.03	0.04	0.01	0.01	0.01	-0.01	-0.01	-0.01	0.01	0.02	0.00	0.02	0.01	-0.00	0.01	0.02	-0.04	-0.01	-0.02	-0.01	0.01	0.01	1.00			
(26) demand	-0.01	-0.02	-0.02	-0.01	0.00	-0.01	0.00	0.06	0.01	-0.02	-0.00	0.02	-0.00	0.02	0.06	-0.02	0.02	0.05	-0.01	0.06	-0.03	-0.00	-0.02	0.01	-0.04	1.00		
(27) price target	0.03	0.00	0.01	-0.01	0.03	-0.02	-0.03	0.01	0.00	-0.03	-0.02	-0.00	-0.01	0.00	-0.01	-0.01	0.00	0.00	0.02	0.01	0.01	0.02	-0.01	0.01	0.00	0.00	1.00	
(28) drilling & pipeline accident	0.01	0.01	0.00	0.00	-0.02	0.00	-0.01	0.01	-0.03	0.01	0.01	-0.01	0.02	0.02	-0.01	0.00	-0.03	0.04	-0.01	0.01	0.01	0.02	-0.02	-0.01	-0.01	-0.01	1.00	

Table A.3

Correlations among the News Indices. This table reports the correlations among the news indices. Data are monthly.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)	(22)	(23)	(24)	(25)	(26)	(27)	(28)
(1) terrorism	1.00																											
(2) war & conflict	−0.00	1.00																										
(3) civil unrest	0.00	−0.16	1.00																									
(4) natural disasters	0.22	−0.08	−0.04	1.00																								
(5) government	0.05	−0.09	0.08	0.01	1.00																							
(6) sovereign debt	0.02	−0.03	−0.08	0.04	0.01	1.00																						
(7) public finance	0.03	−0.10	−0.06	−0.08	0.02	0.20	1.00																					
(8) retail sales	0.22	0.03	−0.10	0.07	0.02	−0.06	0.07	1.00																				
(9) consumer confidence	−0.07	0.18	−0.07	0.14	0.02	−0.06	−0.17	0.01	1.00																			
(10) foreign exchange	0.07	−0.11	0.04	−0.05	−0.12	0.05	−0.02	−0.01	−0.14	1.00																		
(11) housing	0.03	0.13	−0.11	0.05	0.07	−0.01	−0.04	0.10	−0.01	−0.00	1.00																	
(12) interest rates	−0.09	0.04	0.04	−0.05	0.18	−0.05	−0.06	−0.22	−0.05	0.12	0.10	1.00																
(13) treasury yield	−0.08	0.03	−0.02	0.01	0.08	0.02	0.05	−0.10	0.10	−0.19	0.09	0.00	1.00															
(14) private credit	−0.00	−0.24	0.03	−0.06	0.01	0.09	−0.04	−0.11	−0.14	0.06	−0.13	0.01	0.08	1.00														
(15) employment	−0.10	0.13	0.05	0.05	−0.12	−0.17	−0.03	0.20	0.12	−0.03	0.00	−0.09	−0.03	−0.09	1.00													
(16) consumer spending	0.17	−0.06	0.11	0.14	0.11	0.09	0.11	−0.02	0.06	−0.09	−0.01	0.04	−0.03	−0.07	0.07	1.00												
(17) durable goods orders	−0.07	0.07	0.11	0.03	−0.01	−0.07	0.11	0.10	0.15	0.07	0.21	−0.01	0.02	−0.24	0.11	0.00	1.00											
(18) recession	−0.12	−0.04	0.07	−0.08	0.10	−0.13	0.07	0.12	0.05	−0.23	−0.00	−0.23	0.29	0.07	0.09	−0.02	0.14	1.00										
(19) economic growth	0.01	−0.15	−0.09	−0.06	0.00	−0.07	−0.03	0.13	0.01	0.01	−0.02	−0.04	−0.03	0.06	0.30	0.15	0.01	0.10	1.00									
(20) GDP growth	−0.04	−0.07	0.02	−0.11	0.01	−0.04	0.12	0.21	0.03	−0.07	−0.07	−0.23	0.14	0.22	0.20	0.14	0.03	0.50	0.25	1.00								
(21) CPI	0.12	−0.06	0.07	0.05	0.08	−0.03	−0.12	0.05	−0.00	−0.02	−0.02	−0.16	0.08	0.10	0.18	0.13	−0.07	0.20	0.11	0.12	1.00							
(22) PPI	−0.04	−0.16	0.07	−0.03	0.03	0.00	0.16	0.19	−0.01	0.02	−0.05	−0.11	0.18	0.15	0.10	0.08	0.04	0.25	0.17	0.34	0.01	1.00						
(23) exports	−0.21	−0.19	−0.10	−0.18	−0.05	0.08	0.09	−0.09	−0.05	−0.01	−0.20	0.00	0.10	0.17	−0.04	−0.03	−0.01	0.12	0.11	0.29	−0.02	0.14	1.00					
(24) trade balance	−0.13	0.01	−0.12	0.00	0.04	0.05	−0.23	−0.01	−0.08	−0.03	0.10	0.02	−0.03	0.22	−0.06	−0.21	−0.08	0.09	0.00	0.06	−0.00	−0.11	0.08	1.00				
(25) supply	0.08	0.00	−0.01	0.20	0.12	0.13	0.05	−0.12	0.01	−0.06	−0.06	0.02	−0.00	−0.04	−0.35	0.01	−0.03	−0.03	−0.21	0.03	−0.03	−0.01	−0.08	−0.02	1.00			
(26) demand	−0.05	0.08	0.04	−0.06	0.03	−0.23	−0.07	−0.04	0.17	−0.26	0.05	−0.08	0.10	0.14	0.29	0.10	−0.00	0.32	0.18	0.35	0.17	0.16	−0.07	0.11	−0.06	1.00		
(27) price target	−0.01	−0.07	0.09	−0.08	0.08	−0.08	−0.11	0.03	0.03	−0.14	−0.07	−0.05	0.00	0.00	0.12	0.04	−0.08	0.24	0.23	0.12	0.13	0.07	−0.01	0.08	−0.10	0.30	1.00	
(28) drilling & pipeline accident	0.00	−0.05	−0.04	−0.06	−0.03	0.05	−0.15	0.13	0.02	−0.04	−0.09	−0.14	0.07	0.19	−0.10	0.04	−0.11	0.17	−0.02	0.17	0.11	0.04	0.08	0.04	−0.02	−0.00	0.04	1.00

**Table A.4**

Oil returns and news: Controlling for the FX news effect. This table reports the estimates from the regressions of log oil returns on the news indices in each category except for FX news, by using the FX news as an additional control variable in the regressions. Column 1 – 3 report results from the daily data, while column 4–6 report those from the monthly data. The bottom panel reports the control variables used in running the regression. Oil returns are expressed in percentage points. News scores are standardized to have zero mean and unit variance. Newey–West standard errors are used to calculate the *t*-statistical significance. Numbers with \*: statistical significance at 10% level; \*: 5% level; \*\*: 1% level.

	Daily data			Monthly data		
	1	2	3	4	5	6
Terrorism	–12.43*	–12.09*	–12.26*	–0.09	0.02	0.02
War & conflict	–13.04*	–12.80*	–13.01*	–0.30	–0.24	–0.24
Civil unrest	10.46*	9.57*	9.04*	–0.19	–0.82	–0.82
Natural disasters	–12.25*	–12.22*	–14.39*	–1.79*	–1.88*	–1.88*
Government	10.22*	9.05*	10.73*	2.12*	2.06*	2.06*
Sovereign debt	9.12*	8.65*	5.27	–0.29	0.08	0.08
Public finance	13.51*	12.11*	9.73*	1.32*	1.11	1.11
Retail sales	13.97*	14.59*	12.76*	0.61	0.78	0.78
Consumer confidence	4.15	3.29	1.70	0.11	–0.19	–0.19
Housing	–0.54	–0.02	3.51	–0.46	–0.05	–0.05
Interest rates	11.91*	12.20*	6.84*	–0.23	0.50	0.50
Treasury yield	10.99*	9.28*	7.48*	2.15*	1.57*	1.57*
Private credit	11.95*	13.52*	15.51*	0.12	0.35	0.35
Employment	13.78*	13.14*	8.60*	1.18	0.70	0.70
Consumer spending	2.42	1.46	0.77	1.60	0.82	0.82
Durable goods orders	–2.43	–3.04	–3.28	–0.62	–0.60	–0.60
Recession	2.71	0.81	–2.90	2.95*	2.22*	2.22*
Economic growth	–1.55	–1.82	0.26	2.02*	1.78*	1.78*
GDP growth	5.45	2.90	0.10	2.02	1.94*	1.94*
CPI	3.22	3.06	5.18	1.66*	1.45*	1.45*
PPI	–1.64	–2.30	–4.67	–0.04	–0.37	–0.37
Exports	4.06	2.26	4.21	1.10	0.79	0.79
Trade balance	0.25	2.29	3.62	0.43	1.00	1.00
Supply	–9.09*	–9.44*	–6.13	–1.36*	–1.76*	–1.76*
Demand	11.21*	10.00*	11.33*	3.10*	2.48*	2.48*
Price target	8.10*	6.03	6.24	4.43*	4.03*	4.03*
Drilling & pipeline accident	–13.30*	–13.34*	–12.98*	0.80	0.42	0.42
Control variables + monthly dummies		Yes	Yes		Yes	Yes
Without crisis period			Yes			Yes
FX news		Yes	Yes		Yes	Yes

**Table A.5**

Robustness check I: Including neutral news scores. This table reports the results based on the alternative news weighting method which also include the number of neutral news  $news_i = \frac{\sum \eta^p ESS_i^p + \sum \eta^z ESS_i^z + \sum \eta^n ESS_i^n}{\eta^p + \eta^z + \eta^n}$ . News coefficients from the regression of daily log oil returns on the news scores in each category, as described in Eq. (1). The bottom panel reports the control variables used in running each of the regressions. Column 1–5 report the news coefficients with daily data. The last part reports the news coefficients with monthly data. Oil returns are expressed in basis points. News scores are standardized to have zero mean and unit variance. Newey–West standard errors are used to calculate the *t*-statistical significance. Numbers with \*: statistical significance at 10% level; \*: 5% level; \*\*: 1% level.

	Daily data					Monthly data				
	1	2	3	4	5	1	2	3	4	5
Terrorism	–12.43*	–12.42*	–12.72*	–12.08*	–12.27*	–0.09	–0.26	–0.26	–0.09	–0.26
War & conflict	–13.04*	–12.57*	–12.69*	–12.70*	–12.30*	–0.30	0.01	0.01	–0.32	0.01
Civil unrest	8.91*	8.21*	7.47*	9.09*	7.30*	–0.33	–1.06	–1.06	–0.33	–1.05
Natural disasters	–12.25*	–11.86*	–13.65*	–12.62*	–14.22*	–1.79*	–1.85*	–1.85*	–1.79*	–1.85*
Government	9.67*	9.78*	11.90*	9.00*	12.15*	1.98*	2.22*	2.22*	1.75*	1.98*
Sovereign debt	9.12*	9.34*	6.00*	8.19*	4.73	–0.29	–0.14	–0.14	–0.44	–0.28
Public finance	13.51*	12.90*	10.47*	13.69*	10.85*	1.32*	1.06	1.06	1.12	0.88
Retail sales	13.97*	14.21*	12.32*	14.50*	12.58*	0.61	0.82	0.82	0.68	0.89
Consumer confidence	4.15	4.56	2.28	3.45	1.53	0.11	0.31	0.31	0.09	0.25
Foreign exchange	–29.12*	–29.13*	–26.04*	–20.77*	–18.40*	–2.78*	–2.89*	–2.89*	–2.11*	–2.19*
Housing	–0.54	–0.18	3.06	–0.78	2.81	–0.46	–0.12	–0.12	–0.47	–0.09
Interest rates	12.04*	12.57*	7.19*	11.65*	6.46*	–0.13	0.22	0.22	–0.15	0.21
Treasury yield	11.68*	11.11*	10.13*	11.07*	9.25*	2.21*	1.96*	1.96*	2.13*	1.87*
Private credit	11.95*	12.14*	14.73*	11.45*	13.18*	0.12	0.18	0.18	0.14	0.20

(continued on next page)

Table A.5 (continued).

	Daily data					Monthly data				
	1	2	3	4	5	1	2	3	4	5
Employment	13.78*	13.66*	9.10*	13.75*	9.33*	1.18	0.86	0.86	1.18	0.89
Consumer spending	2.42	1.79	0.76	1.34	0.02	1.60	1.19	1.19	1.45	1.01
Durable goods orders	−2.43	−2.74	−3.27	−3.00	−3.55	−0.62	−0.90	−0.90	−0.46	−0.70
Recession	2.71	1.20	−2.57	1.63	−3.37	2.95*	2.83*	2.83*	2.69*	2.56*
Economic growth	−1.55	−2.56	−0.50	−0.82	−0.31	2.02*	1.85*	1.85*	1.91*	1.71*
GDP growth	5.45	4.05	1.29	5.05	1.20	2.02	2.35*	2.35*	1.96	2.33*
CPI	3.22	2.96	5.23	4.18	5.74	1.66*	1.57*	1.57*	1.57*	1.47*
PPI	−1.64	−2.39	−4.58	−1.56	−4.53	−0.04	−0.48	−0.48	−0.02	−0.32
Exports	5.45	4.51	6.70*	4.54	4.96	1.32*	0.80	0.80	1.27*	0.78
Trade balance	0.83	1.53	3.79	0.57	3.84	0.42	1.02	1.02	0.39	1.00
Supply	−9.09*	−9.23*	−5.80	−9.89*	−6.01	−1.36*	−1.54*	−1.54*	−1.35*	−1.45*
Demand	11.21*	10.64*	11.99*	11.14*	11.57*	3.10*	3.16*	3.16*	3.11*	3.20*
Price target	8.10*	7.08*	7.36*	7.99*	6.93*	4.43*	4.46*	4.46*	4.15*	4.13*
Drilling & pipeline accident	−13.30*	−13.65*	−13.45*	−13.34*	−13.52*	0.80	0.63	0.63	0.76	0.55
Control variables + monthly dummies		Yes	Yes		Yes		Yes	Yes		Yes
Without crisis period			Yes		Yes			Yes		Yes
SDR instead of Dollar price				Yes	Yes				Yes	Yes

Table A.6

Robustness check II: Sum up news scores instead of taking the average. This table reports the results based on the alternative news weighting method which sums up the ESS scores of the same day instead of taking the average. News coefficients from the regression of daily log oil returns on the news scores in each category, as described in Eq. (1). The bottom panel reports the control variables used in running each of the regressions. Column 1–5 report the news coefficients with daily data. The last part reports the news coefficients with monthly data. Oil returns are expressed in basis points. News scores are standardized to have zero mean and unit variance. Newey–West standard errors are used to calculate the *t*-statistical significance. Numbers with \*: statistical significance at 10% level; \*: 5% level; \*: 1% level.

	Daily data					Monthly data				
	1	2	3	4	5	1	2	3	4	5
Terrorism	−13.56*	−13.52*	−14.05*	−13.21*	−13.60*	0.02	−0.11	−0.11	−0.01	−0.13
War & conflict	−11.54*	−10.96*	−12.94*	−11.13*	−12.37*	−0.04	0.38	0.38	−0.04	0.41
Civil unrest	5.66	4.84	5.43	5.77	5.37	0.05	−0.56	−0.56	0.09	−0.50
Natural disasters	−8.49*	−8.42*	−9.03*	−8.60*	−9.26*	−0.95	−1.12	−1.12	−0.97	−1.19
Government	7.29*	7.38*	8.07*	6.38*	7.93*	1.76*	1.93*	1.93*	1.58*	1.75*
Sovereign debt	9.29*	9.50*	5.71	8.37*	4.47	−0.43	−0.33	−0.33	−0.58	−0.47
Public finance	16.95*	16.30*	12.97*	16.58*	12.95*	1.25	0.92	0.92	1.11	0.79
Retail sales	11.28*	11.46*	10.33*	11.97*	10.62*	0.75	0.89	0.89	0.79	0.91
Consumer confidence	4.55	4.94	2.91	3.95	2.31	0.12	0.32	0.32	0.15	0.33
Foreign exchange	−34.38*	−34.54*	−25.97*	−24.06*	−16.87*	−3.45*	−3.56*	−3.56*	−2.78*	−2.87*
Housing	0.19	0.38	4.05	−0.03	3.88	0.20	0.47	0.47	0.15	0.47
Interest rates	12.01*	12.79*	7.48*	11.78*	6.88*	−0.30	0.12	0.12	−0.26	0.20
Treasury yield	8.08*	7.86*	7.46*	7.30*	6.45*	1.86*	1.84*	1.84*	1.70*	1.67*
Private credit	12.06*	12.35*	13.68*	11.67*	12.24*	−0.17	−0.04	−0.04	−0.13	−0.02
Employment	11.10*	10.87*	7.99*	11.31*	8.25*	1.10	0.73	0.73	1.13	0.77
Consumer spending	2.13	1.72	2.47	1.11	1.77	1.89*	1.70*	1.70*	1.65*	1.42*
Durable goods orders	−2.11	−2.29	−2.21	−3.11	−2.89	−0.98	−1.17	−1.17	−0.81	−0.95
Recession	0.14	−1.49	−1.55	−0.44	−1.88	3.24*	3.18*	3.18*	3.00*	2.94*
Economic growth	−1.58	−2.67	−0.72	−1.10	−0.77	1.71*	1.41*	1.41*	1.64	1.32*
GDP growth	7.30*	6.14	3.50	6.59	2.97	2.04	2.44*	2.44*	2.01*	2.47*
CPI	2.53	2.09	7.13*	3.00	7.17*	1.94*	2.13*	2.13*	1.90*	2.11*
PPI	−2.22	−3.06	−5.78	−2.32	−5.99	−0.12	−0.66	−0.66	−0.08	−0.51
Exports	8.84*	8.09*	11.61*	8.44*	10.20*	1.11	0.70	0.70	1.13	0.78
Trade balance	0.73	1.33	3.87	0.65	3.87	0.50	1.13	1.13	0.44	1.09
Supply	−8.44*	−8.67*	−4.42	−8.92*	−4.37	−1.46*	−1.78*	−1.78*	−1.44*	−1.67*
Demand	11.21*	10.64*	11.99*	11.14*	11.57*	3.10*	3.16*	3.16*	3.11*	3.20*
Price target	8.44*	7.44*	7.46*	8.28*	6.98*	4.34*	4.35*	4.35*	4.06*	4.02*
Drilling & pipeline accident	−12.30*	−12.53*	−12.27*	−12.35*	−12.34*	0.69	0.60	0.60	0.67	0.55
Control variables + monthly dummies		Yes	Yes		Yes		Yes	Yes		Yes
Without crisis period			Yes		Yes			Yes		Yes
SDR instead of Dollar price				Yes	Yes				Yes	Yes



Table A.7

Robustness check III: Taking only the sign of the news scores. In order to check if our results are sensitive to the magnitude of news scores, we only take the sign of the news scores to construct the news indices, before weighting the news as  $news_i = \frac{\sum_i^p ESS_i^p + \sum_i^z ESS_i^z + \sum_i^n ESS_i^n}{\eta^p + \eta^z + \eta^n}$ . News coefficients from the regression of daily log oil returns on the news scores in each category, as described in Eq. (1). The bottom panel reports the control variables used in running each of the regressions. Column 1–5 report the news coefficients with daily data. The last part reports the news coefficients with monthly data. Oil returns are expressed in basis points. News scores are standardized to have zero mean and unit variance. Newey–West standard errors are used to calculate the  $t$ -statistical significance. Numbers with \*: statistical significance at 10% level; \*: 5% level; \*\*: 1% level.

	Daily data					Monthly data				
	1	2	3	4	5	1	2	3	4	5
Terrorism	−12.24*	−12.14*	−12.18*	−11.84*	−11.65*	−0.17	−0.27	−0.27	−0.20	−0.30
War & conflict	−13.38*	−12.94*	−13.51*	−12.57*	−12.86*	−0.16	0.13	0.13	−0.17	0.13
Civil unrest	9.37*	8.81*	7.83*	9.39*	7.72*	−0.58	−1.18	−1.18	−0.52	−1.08
Natural disasters	−11.10*	−10.56*	−12.32*	−11.53*	−12.96*	−1.73*	−1.77*	−1.77*	−1.76*	−1.80*
Government	8.03*	8.04*	11.44*	7.43*	11.85*	2.21*	2.39*	2.39*	2.02*	2.18*
Sovereign debt	8.74*	8.97*	5.09	7.84*	3.88	−0.31	−0.15	−0.15	−0.46	−0.28
Public finance	11.85*	11.33*	9.62*	12.04*	10.06*	0.87	0.65	0.65	0.71	0.54
Retail sales	11.98*	11.82*	7.43*	12.56*	8.05*	0.98	0.99	0.99	1.20	1.24
Consumer confidence	5.02	5.47	2.91	4.23	2.05	0.02	0.23	0.23	0.01	0.18
Foreign exchange	−29.11*	−29.12*	−26.03*	−20.77*	−18.40*	−2.78*	−2.89*	−2.89*	−2.11*	−2.19*
Housing	−0.61	−0.24	1.88	−1.05	1.43	−0.72	−0.44	−0.44	−0.69	−0.37
Interest rates	11.99*	12.52*	7.19*	11.56*	6.47*	−0.10	0.25	0.25	−0.12	0.24
Treasury yield	11.24*	10.53*	9.27*	10.52*	8.19*	2.32*	2.04*	2.04*	2.24*	1.94*
Private credit	12.66*	12.86*	12.74*	11.85*	11.34*	0.04	0.14	0.14	0.06	0.16
Employment	14.94*	14.75*	11.45*	15.03*	11.75*	1.39	1.05	1.05	1.43	1.12
Consumer spending	1.17	0.65	1.73	−0.04	0.73	1.57	1.39	1.39	1.45	1.25
Durable goods orders	−1.30	−1.68	−3.54	−1.66	−3.58	−0.25	−0.59	−0.59	−0.11	−0.40
Recession	2.71	1.20	−2.57	1.63	−3.37	2.95*	2.83*	2.83*	2.69*	2.56*
Economic growth	−1.49	−2.49	−0.42	−0.75	−0.22	2.02*	1.88*	1.88*	1.91*	1.74*
GDP growth	6.40	5.05	2.38	6.09	2.38	2.10	2.43*	2.43*	2.02	2.38*
CPI	3.48	3.15	5.33	4.54	5.96	1.47*	1.34*	1.34*	1.40*	1.27*
PPI	−1.21	−1.96	−3.84	−1.18	−3.81	0.06	−0.34	−0.34	0.08	−0.19
Exports	4.68	3.85	5.78	3.84	4.08	1.13*	0.62	0.62	1.11*	0.63
Trade balance	1.87	2.63	4.44	1.80	4.81	0.70	1.44*	1.44*	0.68	1.43*
Supply	−8.49*	−8.74*	−5.17	−8.94*	−5.02	−1.57*	−2.00*	−2.00*	−1.55*	−1.88*
Demand	10.13*	9.74*	10.85*	10.08*	10.45*	3.10*	3.16*	3.16*	3.11*	3.20*
Price target	8.64*	7.61*	8.46*	8.40*	7.90*	4.44*	4.47*	4.47*	4.17*	4.14*
Drilling & pipeline accident	−13.26*	−13.58*	−13.31*	−13.40*	−13.47*	0.93	0.77	0.77	0.90	0.70
Control variables + monthly dummies		Yes	Yes		Yes		Yes	Yes		Yes
Without crisis period			Yes		Yes			Yes		Yes
SDR instead of Dollar price				Yes	Yes				Yes	Yes

Table A.8

Monthly predictive regressions: Controlling for crisis period and with control variables. This table reports the results from the regression Eq. (3) after excluding the crisis period. Newey–West standard errors are used to calculate the  $t$ -statistical significance. Numbers in italics: statistical significance at 10% level; bold: 5% level; italics and bold: 1% level.

	$news_{t-1}$	$news_{t-2}$	$news_{t-3}$	$news_{t-4}$	$news_{t-5}$	$news_{t-6}$	$adj. R^2$	$\sum_{L=1}^5 \beta(L)$	Wald(6)	$\sum_{L=1}^2 \beta(L)$	Wald(2)
Terrorism	−0.50	−1.08	0.13	−1.14	2.30*	−0.96	0.11	−1.24	0.35	−1.58	0.55
War & conflict	−0.31	0.56	0.43	−1.32	0.51	1.95*	0.12	1.82	0.79	0.25	0.80
Civil unrest	−1.15	−0.90	−0.20	0.01	−1.03	0.77	0.08	−2.50	0.99	−2.05	0.40
Natural disasters	0.26	−1.09	1.90*	−1.86*	0.48	0.17	0.09	−0.14	0.69	−0.83	0.46
Government	−2.05*	1.95*	0.89	−0.14	1.57*	−0.60	0.17	1.61	0.76	−0.11	0.13
Sovereign debt	2.40*	−1.00	−0.85	1.14	2.10*	−3.26*	0.24	0.52	0.25	1.40*	0.03
Public finance	−0.48	−0.83	1.79*	−1.92*	0.21	−0.03	0.10	−1.27	0.53	−1.31	0.77
Retail sales	1.39	−0.24	0.17	−1.69*	−2.16*	0.82	0.13	−1.71	0.50	1.15	0.42
Consumer confidence	−1.12	0.55	0.00	0.29	−0.38	1.61	0.08	0.95	0.99	−0.57	0.72
Foreign exchange	−1.81*	0.26	0.78	1.53*	0.10	0.67	0.10	1.53	0.85	−1.55	0.27
Housing	−0.85	0.43	−0.32	0.01	−1.44*	1.03	0.08	−1.14	0.98	−0.42	0.70
Interest rates	−0.30	−0.61	0.36	−1.14	0.09	−0.54	0.07	−2.14	0.96	−0.91	0.82
Treasury yield	−1.24	0.41	0.53	−2.12*	−0.67	1.02	0.10	−2.07	0.87	−0.83	0.45
Private credit	1.00	0.16	−0.84	2.49*	−0.16	−1.07	0.13	1.59	0.38	1.17	0.66
Employment	0.79	−0.82	2.21*	−0.46	−0.11	−0.42	0.10	1.18	0.60	−0.03	0.80
Consumer spending	0.55	−1.22*	0.98	−1.56*	0.30	0.19	0.10	−0.77	0.97	−0.67	0.41
Durable goods orders	1.19*	−0.79	−0.62	0.43	−0.72	0.75	0.09	0.24	0.99	0.40	0.51
Recession	2.77*	−0.75	0.47	−0.68	−0.99	1.38	0.12	2.19	0.27	2.02	0.12
Economic growth	1.85*	−0.39	0.92	1.63	−0.56	−0.50	0.13	2.96	0.59	1.46	0.23
GDP growth	2.90*	−0.07	−0.46	−2.78*	3.30*	−0.09	0.17	2.79	0.14	2.82	0.13

(continued on next page)

Table A.8 (continued).

	$news_{t-1}$	$news_{t-2}$	$news_{t-3}$	$news_{t-4}$	$news_{t-5}$	$news_{t-6}$	$adj. R^2$	$\sum_{L=1}^5 \beta(L)$	Wald(6)	$\sum_{L=1}^2 \beta(L)$	Wald(2)
CPI	2.86*	-2.37*	1.13	1.11	-0.05	0.15	0.19	2.83	0.40	0.49*	0.00
PPI	0.55	3.06*	2.02	-0.38	-2.74*	0.21	0.16	2.72	0.70	3.61	0.24
Exports	0.66	-0.33	0.74	-0.51	-0.68	0.21	0.05	0.08	0.98	0.33	0.79
Trade balance	0.25	-1.04	0.69	-1.50	0.21	-0.75	0.10	-2.12	0.85	-0.78	0.68
Supply	0.19	-0.51	-0.66	0.58	0.81	-0.74	0.05	-0.33	0.93	-0.32	0.92
Demand	-0.48	0.81	0.36	2.31*	0.99	0.10	0.15	4.09	0.71	0.33	0.62
Price target	-0.46	1.56*	-1.19	0.77	-1.33	-0.15	0.07	-0.79	0.88	1.10	0.23
Drilling & pipeline accident	-0.28	0.65	-0.49	1.20	-0.74	1.62*	0.10	1.97	0.66	0.37	0.77

Table A.9

Net commercial traders positions and news. This table reports the news impact on the net long commercial traders positions as in Eq. (6). The coefficient  $\beta$ ,  $R^2$ , the sum of the coefficients of the five news lags  $\sum_{L=1}^5 \beta(L)$ , and the  $p$ -value of the Wald-test are reported. Newey–West standard errors are used to calculate the  $t$ -statistical significance. The net long commercial traders position is calculated according to Eq. (5) and is expressed in percentage points. The data are weekly. Numbers in italics: statistical significance at 10% level; bold: 5% level; italics and bold: 1% level.

	$news_{t-1}$	$news_{t-2}$	$news_{t-3}$	$news_{t-4}$	$adj. R^2$	$\sum_{L=1}^4 \beta(L)$	Wald(4)	$\sum_{L=1}^2 \beta(L)$	Wald(2)
Terrorism	-0.23*	-0.07	0.01	0.08	0.94	-0.20	0.42	-0.30*	0.10
War & conflict	-0.11*	-0.06	-0.02	-0.01	0.94	-0.20	0.86	-0.17	0.36
Civil unrest	-0.02	0.02	-0.00	0.22*	0.94	0.22	0.56	0.00	0.94
Natural disasters	0.08	-0.03	0.04	-0.14*	0.94	-0.06	0.28	0.04	0.42
Government	0.05	-0.07	0.01	-0.09	0.94	-0.09	0.86	-0.01	0.55
Sovereign debt	0.07	-0.02	0.02	0.03	0.94	0.09	0.96	0.05	0.59
Public finance	-0.01	-0.02	0.01	-0.03	0.94	-0.05	1.00	-0.03	0.97
Retail sales	-0.11*	-0.02	-0.07	0.05	0.94	-0.15	0.90	-0.13	0.46
Consumer confidence	-0.09	-0.05	-0.14*	0.12*	0.94	-0.15	0.53	-0.14	0.49
Foreign exchange	0.02	0.04	0.01	-0.05	0.94	0.03	0.98	0.06	0.84
Housing	-0.08	0.04	-0.05	-0.09*	0.94	-0.17	0.78	-0.04	0.47
Interest rates	-0.03	0.07	-0.03	-0.12*	0.94	-0.11	0.53	0.04	0.45
Treasury yield	-0.05	-0.04	-0.03	0.01	0.94	-0.11	0.97	-0.09	0.72
Private credit	0.13*	-0.03	0.11*	-0.03	0.94	0.18	0.85	0.09	0.49
Employment	-0.11	-0.13*	-0.10	-0.05	0.94	-0.38	0.68	-0.24	0.33
Consumer spending	-0.09	0.04	-0.08	0.05	0.94	-0.07	0.88	-0.05	0.52
Durable goods orders	-0.00	-0.14*	-0.05	0.01	0.94	-0.19	0.84	-0.14	0.29
Recession	-0.10*	-0.02	-0.03	-0.02	0.94	-0.17	0.89	-0.12	0.42
Economic growth	-0.07	-0.13*	0.01	-0.08	0.94	-0.27	0.60	-0.20	0.23
GDP growth	0.07	-0.08	0.03	-0.11*	0.94	-0.10	0.46	-0.01	0.46
CPI	-0.13*	0.02	0.09	0.02	0.94	0.00	0.85	-0.11	0.42
PPI	0.06	0.11	-0.00	-0.02	0.94	0.15	0.93	0.17	0.39
Exports	-0.01	-0.03	0.03	-0.04	0.94	-0.04	0.99	-0.04	0.93
Trade balance	-0.06	0.06	0.04	-0.03	0.94	0.01	0.94	0.00	0.69
Supply	0.05	0.08	0.05	-0.07	0.94	0.11	0.85	0.13	0.46
Demand	-0.07	-0.13*	0.02	0.03	0.94	-0.15	0.59	-0.20*	0.07
Price target	-0.08	0.05	0.04	0.16*	0.94	0.18	0.62	-0.03	0.52
Drilling & pipeline accident	-0.09*	-0.06	0.06	0.15*	0.94	0.06	0.11	-0.15	0.22

## References

- Antweiler, W., Frank, M.Z., 2004. Is all that talk just noise? The information content of internet stock message boards. *J. Finance* 59 (3), 1259–1293.
- Bakshi, G., Kapadia, N., Madan, D., 2003. Stock return characteristics, skew laws, and the differential pricing of individual equity options. *Rev. Financ. Stud.* 16 (1), 101–143.
- Baumeister, C., Kilian, L., 2012. Real-time forecasts of the real price of oil. *Rev. Econ. Stat.* 30 (2), 326–336.
- Baumeister, C., Kilian, L., 2014. A general approach to recovering market expectations from futures prices with an application to crude oil. Working Paper.
- Baumeister, C., Kilian, L., 2015. Forecasting the real price of oil in a changing world: A forecast combination approach. *Rev. Econ. Stat.* 33 (3), 338–351.
- Bernile, G., Hu, J., Tang, Y., 2016. Can information be locked up? informed trading ahead of macro-news announcements. *J. Financ. Econ.* 121 (3), 496–520.
- Bessembinder, H., Chan, K., 1992. Time-varying risk premia and forecastable returns in futures markets. *J. Financ. Econ.* 2 (3), 169–193.
- Bollerslev, T., Tauchen, G., Zhou, H., 2009. Expected stock returns and variance risk premia. *Rev. Financ. Stud.* 22 (3), 4463–4492.
- Boudoukh, J., Feldman, R., Kogan, S., Richardson, M., 2013. Which news moves stock prices? A textual analysis. NBER Working Paper.
- Breunig, R.V., Chia, T.C., 2013. Sovereign ratings and oil-exporting countries: The effect of high oil prices on ratings. Working Paper.
- Chan, W.S., 2003. Stock price reaction to news and no-news: Drift and reversal after headlines. *J. Financ. Econ.* 70 (2), 223–260.
- Cochrane, J.H., Piazzesi, M., 2005. Bond risk premia. *Amer. Econ. Rev.* 95 (1), 138–160.
- Coval, J.D., Shumway, T., 2001. Is sound just noise? *J. Finance* 56 (5), 1887–1910.
- Dang, T.L., Moshirian, F., Zhang, B., 2015. Commonality in news around the world. *J. Financ. Econ.* 116, 82–110.
- Das, S.R., Chen, M.Y., 2007. Yahoo! for amazon: Sentiment extraction from small talk on the web. *Manage. Sci.* 53 (9), 1375–1388.
- Dougal, C., Engelberg, J., Garcia, D., Parsons, C., 2012. Journalists and the stock markets. *Rev. Financ. Stud.* 25 (3), 639–679.
- Dzieliński, M., 2012. Which news resolves asymmetric information? Working Paper.
- Dzieliński, M., Hasseltoft, H., 2015. News tone dispersion and investor disagreement. Working Paper.
- Elder, J., Miao, H., Ramchander, S., 2013. Jumps in oil prices: The role of economic news. *Energy J.* 34 (3).
- Fang, L., Peress, J., 2009. Media coverage and the cross-section of stock returns. *J. Finance* 64, 2023–2052.

- Garcia, D., 2013. Sentiment during recessions. *J. Finance* 68 (3), 1267–1300.
- Gomes, J.F., Kogan, L., Yogo, M., 2009. Durability of output and expected stock returns. *J. Political Econ.* 117 (5), 941–986.
- Hamilton, J., 2008. Understanding crude oil prices. NBER Working Paper.
- Hamilton, J., Wu, C., 2014. Risk premia in crude oil futures prices. *J. Int. Money Finance* 42, 9–37.
- Heath, D., 2016. Macroeconomic factors in oil futures markets. Working Paper.
- Hitzemann, S., 2016. Macroeconomic fluctuations, oil supply shocks, and equilibrium oil futures prices. Working Paper.
- Hong, H., Stein, J., 1999. A unified theory of underreaction, momentum trading and overreaction in asset markets. *J. Finance* 54 (6), 2143–2184.
- Kilian, L., 2009. Not all oil price shocks are alike: Disentangling demand and supply shocks in the crude oil market. *Amer. Econ. Rev.* 99 (3), 1053–1069.
- Kilian, L., Murphy, D.P., 2014. The role of inventories and speculative trading in the global market for crude oil. *J. Appl. Econometrics* 29, 454–478.
- Kilian, L., Vega, C., 2011. Do energy prices respond to U.S. macro news? a test of the hypothesis of predetermined energy prices. *Rev. Econ. Stat.* 93 (2).
- Loughran, T., McDonald, B., 2011. When is a liability not a liability? Textual analysis, dictionaries, and 10-ks. *J. Finance* 66 (1), 35–65.
- Peress, J., 2014. The media and the diffusion of information in financial markets: Evidence from newspaper strikes. *J. Finance* 69 (5), 2007–2043.
- Petersen, M.A., 2004. Information: Hard and soft. Working Paper.
- Rapport, A.A., 2013. Supply and demand shocks in the oil market and their predictive power. Working Paper.
- Ready, R., 2013. Commodity trade and the carry trade: A tale of two countries. NBER Working Paper.
- de Roon, F.A., Nijman, T.E., Veld, C., 2000. Hedging pressure effects in futures markets. *J. Finance* 55 (3), 1437–1456.
- Smales, L.A., 2014. News sentiment in the gold futures market. *J. Bank. Finance* 49, 275–286.
- Smales, L.A., 2015. Asymmetric volatility response to news sentiment in gold futures. *J. Int. Financ. Markets, Inst. Money* 34, 161–172.
- Soo, C.K., 2015. Quantifying animal spirits: News media and sentiment in the housing market. Working Paper.
- Tetlock, P.C., 2007. Giving content to investor sentiment: The role of media in the stock market. *J. Finance* 62 (3), 1139–1168.
- Tetlock, P.C., 2010. Does public financial news resolve asymmetric information?. *Rev. Financ. Stud.* 23 (9), 3520–3557.
- Tetlock, P.C., 2011. All the news that's fit to reprint: Do investors react to stale information. *Rev. Financ. Stud.* 24 (5), 1481–1512.
- Tetlock, P.C., Saar-Tsechansky, M., Macskassy, S., 2008. More than words, quantifying language to measure firms' fundamentals. *J. Finance* 63 (3), 1437–1467.
- Uhl, M.W., 2014. Reuters sentiment and stock returns. *J. Behav. Finance* 4, 287–298.
- Zhang, S.S., 2013. Need for speed: An empirical analysis of hard and soft information in a high frequency world. Working Paper.