Macro Fundamentals or Geopolitical Events? A Textual Analysis of News Events for Crude Oil

Michael W. Brandt*
Duke University and NBER

Lin Gao†
University of Luxembourg

August 2016

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 not predictability. Moreover, geopolitical news generate more uncertainty and greater trading volume, consistent with a disagreement explanation, while macroeconomic news reduce informational asymmetry and are associated with subsequent lower trading volume. Finally, we find that news sentiment contains more information about future expectations than about future realizations of economic data.

JEL classification: G12, G13.

Keywords: Crude oil, news analytics, sentiment, information

*Fuqua School of Business, Duke University. Address: 100 Fuqua Drive, Box 9012, Durham, NC 27708, USA. E-mail: mbrandt@duke.edu. Tel.: +1 919 660 1948.

†Luxembourg School of Finance, University of Luxembourg. Address: 4, rue Albert Borschette, L-1246 Luxembourg. E-mail: lin.gao@uni.lu. Tel.: +352 466 644 5482.

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. Kilian and Baumeister (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 is 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 provide 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.

We consider a broad cross-section of macroeconomic and geopolitical news. The goal of the 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 as well as more subtly differences between news from oil-producing versus oil-consuming countries. News about economic growth is the strongest predictor of oil returns over coming months, confirming the findings of Kilian (2009) that macroeconomic growth generates demand for oil. The gradual diffusion of public information and resulting momentum and predictability in oil returns can 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). The response of oil prices to geopolitical events is even stronger and immediate. News is incorporated into the oil price at once without follow-on momentum, consistent with informational efficiency.

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 predicts 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, in the case of crude oil, that geopolitical news such as terrorism, war and conflict, civil unrest, and natural

disasters cause a statistically significant increase in trading volume for the next two to three days. Following this period of more active trading, we observe a distinct reversal of volume for the rest of the week (between the 4th and 5th day). This pattern suggests that these geopolitical events cause investors to disagree about the future oil price trajectory. 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 these public releases diminish the informational asymmetry. 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.

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.

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. 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 (Boudoukh et al., 2013). Soo (2015) studies an important single market like us (the housing market in her case) but she focuses on only housing related news sentiment. 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.

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 primarily processed by non high-frequency traders and 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.

Finally, our study is related to the literature on oil variance decomposition. 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. As with any econometric model, these findings 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. RavenPack 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). In addition, RavenPack also assigns each news a novelty score and a relevance score. A novelty score of 100 (ranging from one 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 tone. 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. Our sample ranges from the inception of the dataset is January 1, 2000 through March 31, 2013.

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[1], 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. Economic data releases are surveys or official releases of economic

[1] `http://www.cmegroup.com/market-data/files/cme-group-settlement-procedures.pdf`

statistics. Surrounding either type of release, opinions and comments appear in the media, which may predict the possible outcome or emphasize the importance of the information.

For example, war and conflicts are usually event driven, whereas 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 and lagged news scores at the daily level, whereas for opinions and comments, significance is found forward looking and at the monthly horizon.

The second important aspect of the data is that for some categories there exists an important difference between oil-producing and oil-consuming countries. For example, the exchange rates from countries of commodity currencies typically move conversely to countries of funding currencies in a carry trade. The latter happen to be among the largest oil-consuming countries as well (see, for example, Ready (2013)). The same goes for exports and trade balance. Civil unrest in 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, 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 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).

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 Libya 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).

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[2]. 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 29 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.[3] Also, as already mentioned, within some categories, we find opposing signs between oil-producing and oil-consuming countries. For example, civil unrest in 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 oil-consuming countries, we subtract the score of the oil-producing countries from those of the oil-consuming countries before averaging.

Table 1 describes the data. It provides summary statistics on the 29 news indices as well crude oil returns. The standard deviation of the news scores ranges from zero (for recession indicators) to 38.90. Some categories such as terrorism and civil unrest exhibit predominantly negative scores. In total, 84306 news items are included in our sample. The sample mean is 1.30, which is close to neutral. The standard deviation of the whole sample is as large as 33. We standardize these scores to work with zero mean and unit standard deviation regressors in the empirical analysis below.

[2]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.

[3]We ignore neutral news as it is deemed uninformative one way or the other.

4. Empirical Results

4.1. Regression Analysis and Variance Decomposition

The 29 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.

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.

If the market is efficient, the information should be incorporated into prices immediately, especially when the news is publicly released and broadly disseminated. Therefore, we expect that prices change immediately after the news release. To test this, we regress the log oil returns ΔOil_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[4], 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:

$$\Delta Oil_t = \alpha + \beta \cdot News_{i,t} + \sum_j \varphi_j \cdot Control_{j,t} + M_t + \varepsilon_t \tag{1}$$

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 controlls. 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.

[4]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.

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 14.41 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 affect oil returns negatively both at the daily and monthly level. Raising the nominal interest rates by one standard deviation is expected to slow economic growth (hence a negative ESS) but it pushes up the daily oil return by 9.19 basis points and the monthly return by 1.41 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 more sensitive to demand news than supply news. A negative supply shock of one-standard deviation moves the oil returns up by 8.80 basis points (p-value$<$0.05) at the daily level and 1.73 percentage points at the monthly level (p-value$<$0.1), whereas a positive demand shock raises oil returns by 11.65 basis points (p-value$<$0.01) at the daily level and 3.01 percentage points at the monthly level (p-value$<$0.01). This evidence 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)[5]. The fourth column of Table 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

[5] `http://www.imf.org/external/np/fin/data/rms_sdrv.aspx`

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 figures 3 and 4, 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 sample. Specifically, we adopt the time-line of the financial crisis established by the Federal Reserve Bank of St. Louis[6]. 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.

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 25 categories (geo + macro) for the daily data. The adjusted R^2 of this regression is 6.66 percent. We then regress oil returns on the geopolitical and macroeconomic indexes separately. The adjusted R^2 are 1.85 and 4.81 percent, respectively, which correspond to 28 and 72 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 news increases (74.17%) while that of geopolitical news falls slightly (25.83%), meaning that at the monthly level macroeconomic news plays a more important role. As we discovered previously, the effect of geopolitical news is more transient and tends to wash out over longer horizons.

[6] `http://www.imf.org/external/np/fin/data/rms_sdrv.aspx`

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 next test for similar predictability of oil returns using lagged macroeconomic and geopolitical news. 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:

$$\Delta Oil_t = \alpha + \sum_{k=1}^{K} \beta_k L^k(News_{i,t}) + \sum_{k=1}^{K} \gamma_k L^k(\Delta Oil_t) + \sum_{k=1}^{K} \varphi_k L^k(Control_t) + M_t + \varepsilon_t. \quad (2)$$

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.

At the daily level in Table 6, significant predictive power is found for the first lag of the war and conflict, interest rate, treasury yield and recession categories. Their signs of the coefficients are the same as with those in the contemporaneous regressions, suggesting momentum as well as slow digestion of information. The first two lags of civil unrest and natural disasters are negative, and their magnitude is large in economic sense, consistent with event-day overreaction to this type of geopolitical news. Toward the end of the week, days 4 and 5, we then observe mean reversion (i.e., coefficients that have the opposite sign as the contemporaneous ones), in line with the results of Tetlock (2007). All of these results are robust to excluding the crisis periods (see Table A.1).

At the monthly frequency in Table 7, 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=2.26) over the subsequent month, which is both statistically significant at the five percent level and the scale of the coefficient is economically large. After excluding the financial crisis period in Table A.2, the predictive ability of GDP news becomes even stronger. The coefficient of the first news lag increases both in magnitude and statistical significance (coef=4.36, t-stat=3.16). The Wald-test shows that the sum of the first two lags and the sum of the first six lags are statistically significant at one and 10 percent levels, respectively. Other macroeconomic news categories such as recession, economic growth, CPI, and exports predict oil returns significantly at the first lag. In a Wald-test of the null hypothesis that the six lags of the news index do not forecast returns, we can reject the null for GDP growth,

exports and private credit. It is worth mentioning that the PPI news index leads the CPI index and therefore predict oil returns two to three months in advance. 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.

To illustrate the impact of the size and duration of these predictive news effects, we conduct event studies and present the cumulative oil returns from 20 days before the news arrival to 20 days after. Figure 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. Figure 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 6, 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. However, the supply increase causes higher price volatility and uncertainty. 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 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.

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. Figure 3 and 4 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. 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.

4.4. Trading Volume and News Sentiment

Traders in the market read and digest difffernt quantities of news every day, which is then reflected in variations in trading volume over time. Theories predict contradictory signs of how the media impacts trading volume. 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 volatility and transaction costs, which in turn leads to a reduction in trading activity. On the contrary, 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).

We formally test these types of hypotheses for the news categories in our sample. In the contemporaneous regressions, we apply two kinds of news scores to measure the news impact on the trading volume. First, we use the original news scores adopted before. Second,

we take the absolute value of the news score in our regressions. 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. Table 8 reports the results. Interestingly, consumer confidence, treasury yield, and crude oil demand news are all associated with a significant increase in trading volume. The raw sentiment measure does not cause any volume difference or is not as significant as the absolute one, suggesting that news from both categories triggers disagreement among investors. The coefficients for foreign exchange, consumer spending, and CPI news are negative and statistically significant, suggesting that the release of this information reduces the informational asymmetry among traders and hence decreases trading volume. The extreme news scores in the geopolitical news categories tend to generate volume increases, whereas those in the macroeconomic categories do not.

Next, we estimate and 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:

$$\Delta Vlm_t = \alpha + \sum_{k=1}^{K} \beta_k L^k(|News_{i,t}|) + \sum_{k=1}^{K} \gamma_k L^k(Vlm_t) + \varepsilon_t. \tag{3}$$

Table 9 reports the results. Geopolitical news such as terrorism, war and conflict, civil unrest, and natural disasters causes a statistically significant increase in oil trading volume over the next two to three days, whereas at a later stage (between the 4th and 5th day) a distinct mean reversion trading activity takes place. For instance, a one-standard deviation increase in natural disaster news leads to a growth in trading volume by 1.96 percentage points in the first two days, which is large economically. These results suggest that investor disagree about the future oil price development, which leads to more trading activity. In contrast, macroeconomic news such as consumer spending, durable goods orders, housing, economic growth, CPI and exports dampens trading volume after news release. This may be attributed to the fact that these types of news diminish the information asymmetry and hence enhance market efficiency.

4.5. 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). Here, similar to de Roon et al. (2000)

we define net long positions of a particular trader group i as:

$$NP_{i,t} = \frac{\sharp LongPositions_{i,t} - \sharp ShortPositions_{i,t}}{\sharp LongPositions_{i,t} + \sharp ShortPositions_{i,t}}. \quad (4)$$

The two main trader groups are commercial traders, who mainly use futures contracts for hedging purposes, and non-commercial traders who 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.

Figure 5 plots the net long positions of non-commercial traders and commercial traders calculated as in equation (4). 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), 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 Appendix Table A.3. The table reports results for the following regression specification:

$$NPi,t = \alpha + \sum_{k=1}^{K} \beta_k L^k(News_{i,t}) + \sum_{k=1}^{K} \gamma_k L^k(NP_{i,t}) + \sum_{k=1}^{K} \varphi_k L^k(Control_t) + M_t + \varepsilon_t, \quad (5)$$

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 0.96 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 respond to growing demand for oil with increasing long positions, whereas commercial traders increase their net short positions (decrease long positions). Among the important macroeconomic categories, a one standard deviation increase of good news from the housing market leads to a growth in non-commercial net positions by 0.74 percentage points. News on interest rate raises induce non-commercial traders to increase their net positions. 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.

4.6. 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, 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 equation (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. For instance, 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. In particular, 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. Good news decreases investor's risk aversion, while bad news causes fear of future uncertainty.

4.7. 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). As a motivating example, Figure 6 plots the news sentiment index for durable goods orders and the actual released data. Both time series have been standardized to

ease comparison. The news index 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. The recovery speed in the aftermath of the financial crisis is slow for the news index. It moves sideways between 2009 and 2011, while the actual data pick up the increasing trend rapidly. This evidence suggests anecdotally that the news sentiment index for durable goods 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. This evidence supports the findings in Gomes et al. (2009) who 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.

We provide more formal evidence by regressing the historical durable goods orders data dur_t^{hist} on its own lag and the first lag of the respective news index dur_{t-1}^{news}. To avoid the case that the news data simply reflect the data release, we exclude the news data on the press release days. We keep both data at the monthly frequency. The dependent variable dur_t^{hist} is strongly correlated with dur_{t-1}^{news} (t-stat=3.57, $adj.R^2$=0.88), suggesting there is indeed predictive power of the news index for the official data release.

In Figure 7 we plot as another example the news sentiment index for crude oil supply against the actual released data of US ending stocks excluding SPR of crude oil. The news index tracks the behavior of the released data well although it is more volatile than the actual data. It appears that in this case too 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 correctly and timely.

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 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 not predictability. Moreover, geopolitical news generate more uncertainty and greater trading volume, consistent with a disagreement explanation, while

macroeconomic news reduce informational asymmetry and are associated with subsequent lower trading volume. Finally, we find that news sentiment contains more information about future expectations than about future realizations of economic data.

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.

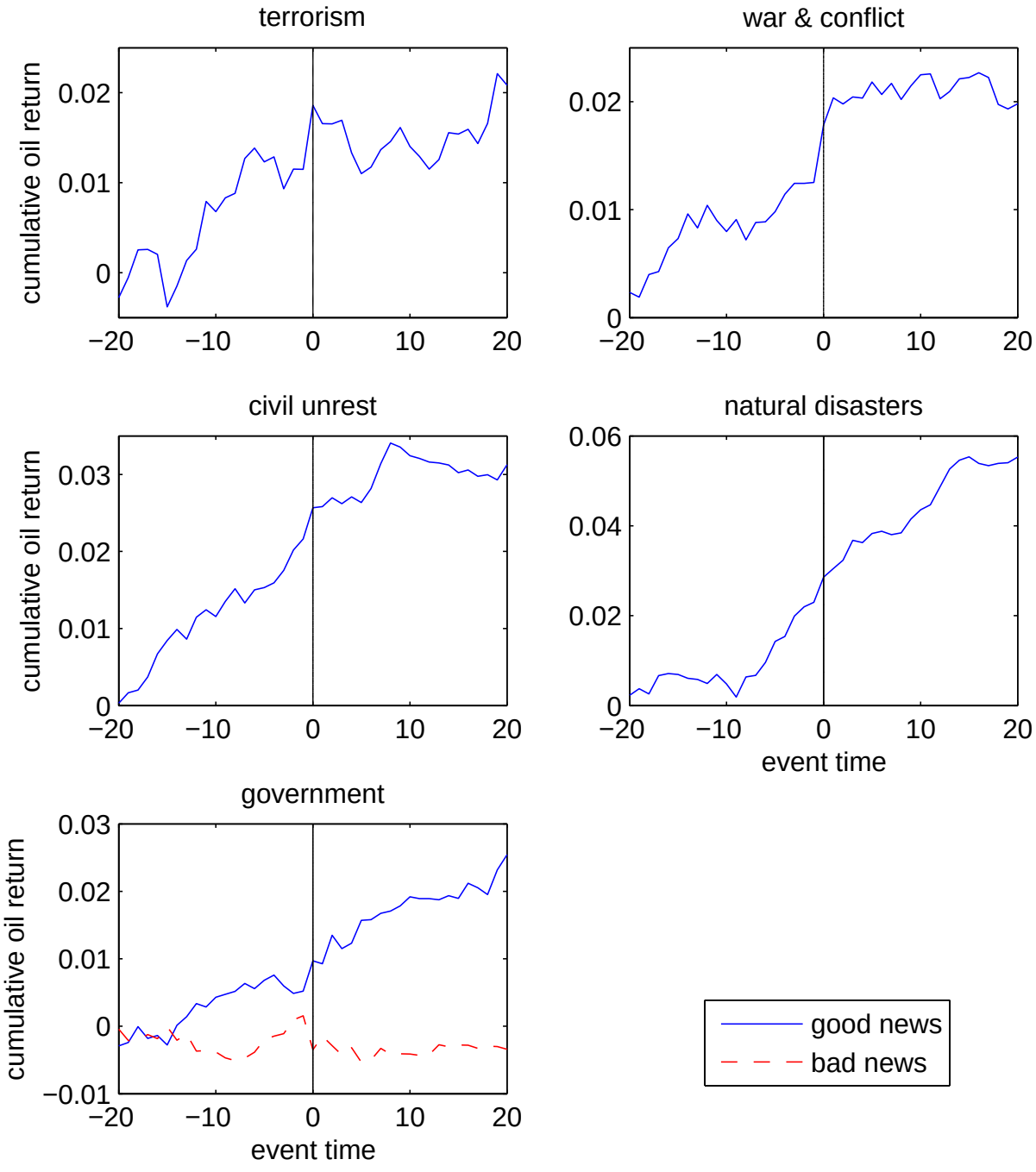

Fig. 2. Event Studies: Economic Impact

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.

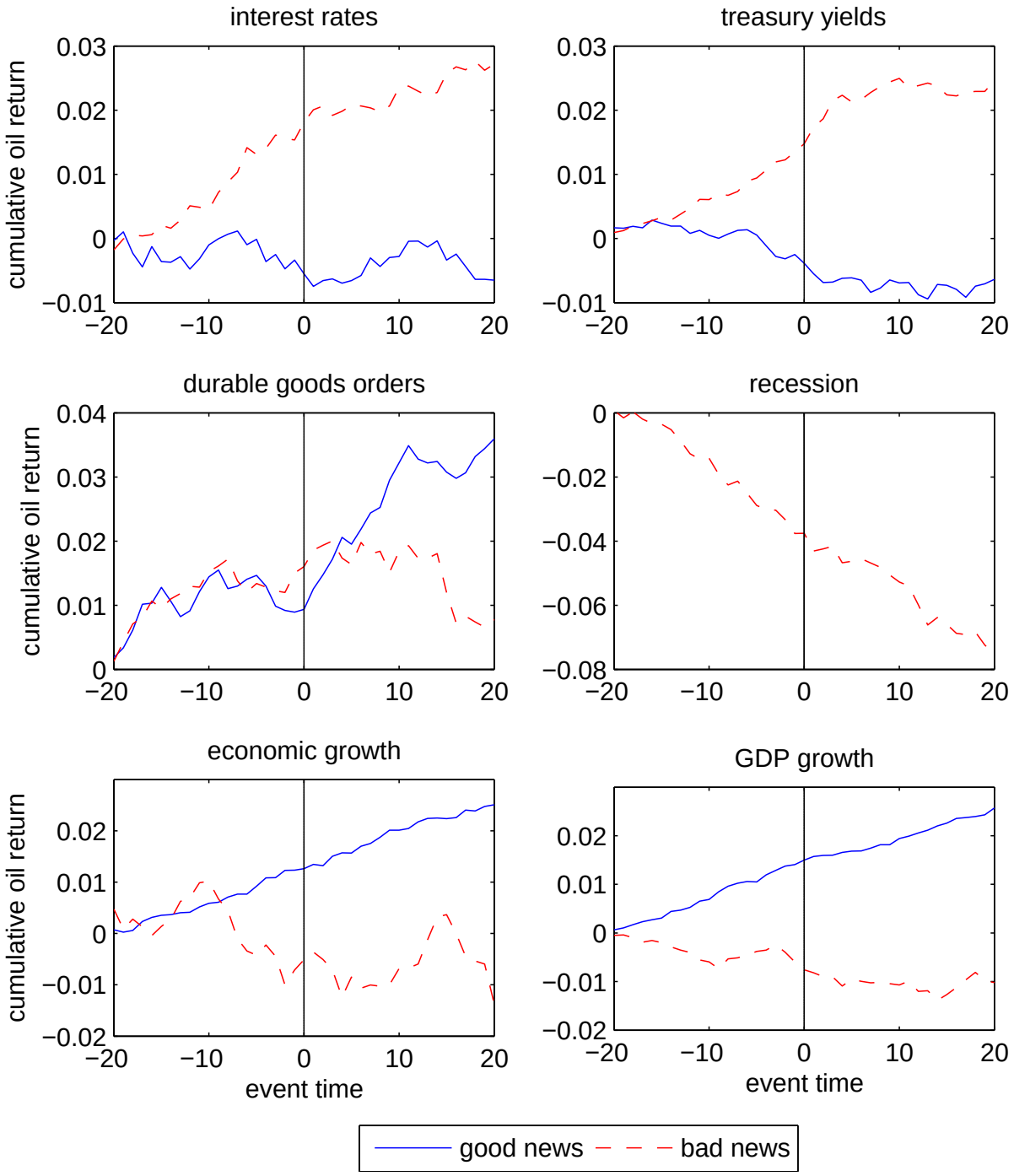

Fig. 2 (Cont.). Event Studies: Economic Impact

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.

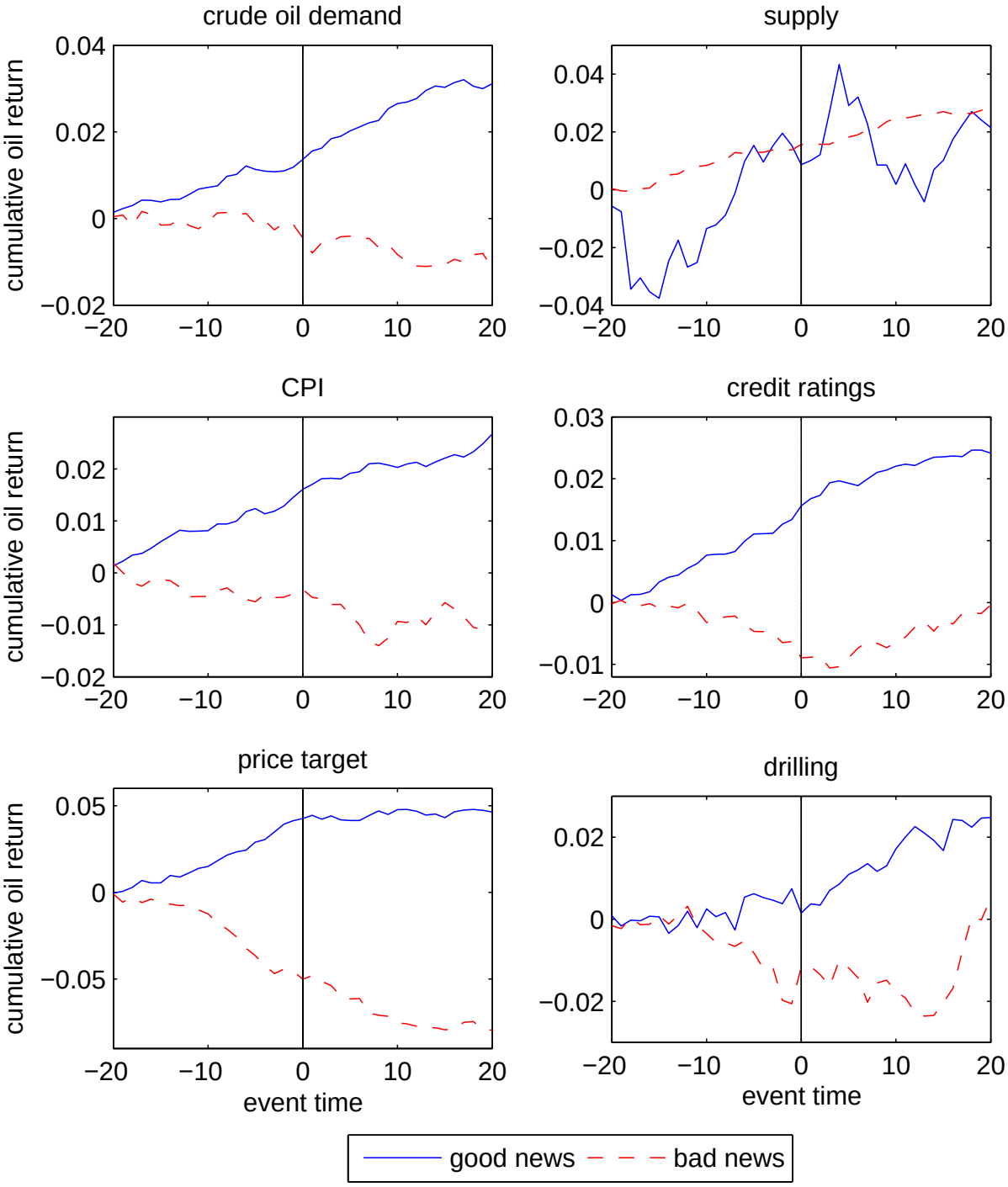

Fig. 3. Geopolitical News Index and the Crude Oil Prices

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.

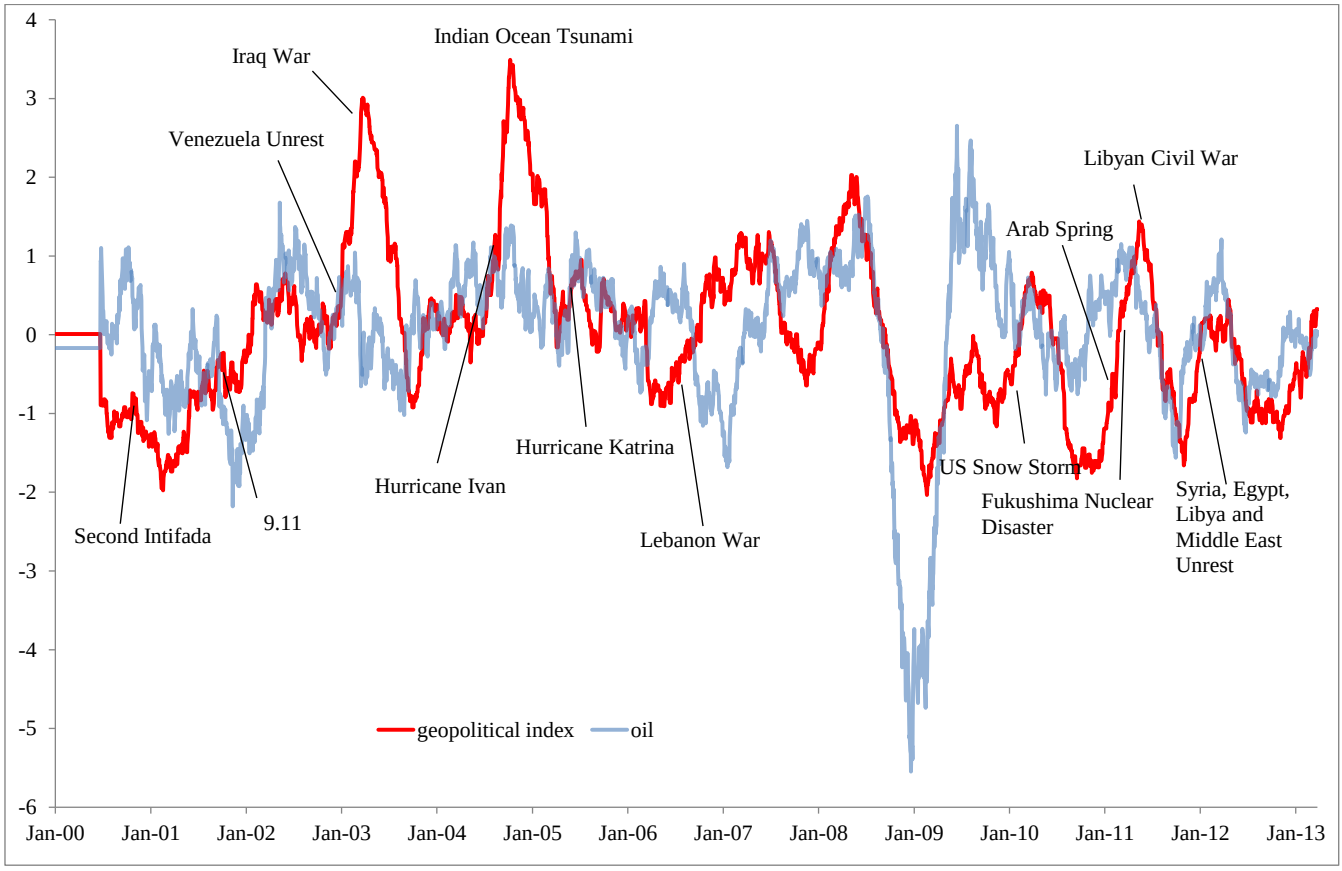

Fig. 4. Aggregate Macro News Index and the Crude Oil Prices

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. 5. Net Positions of Commercial and Noncommerical Traders

This figure coplots the net positions of commercial and noncommercial traders of the crude oil futures contracts calculated according to Equation 4.

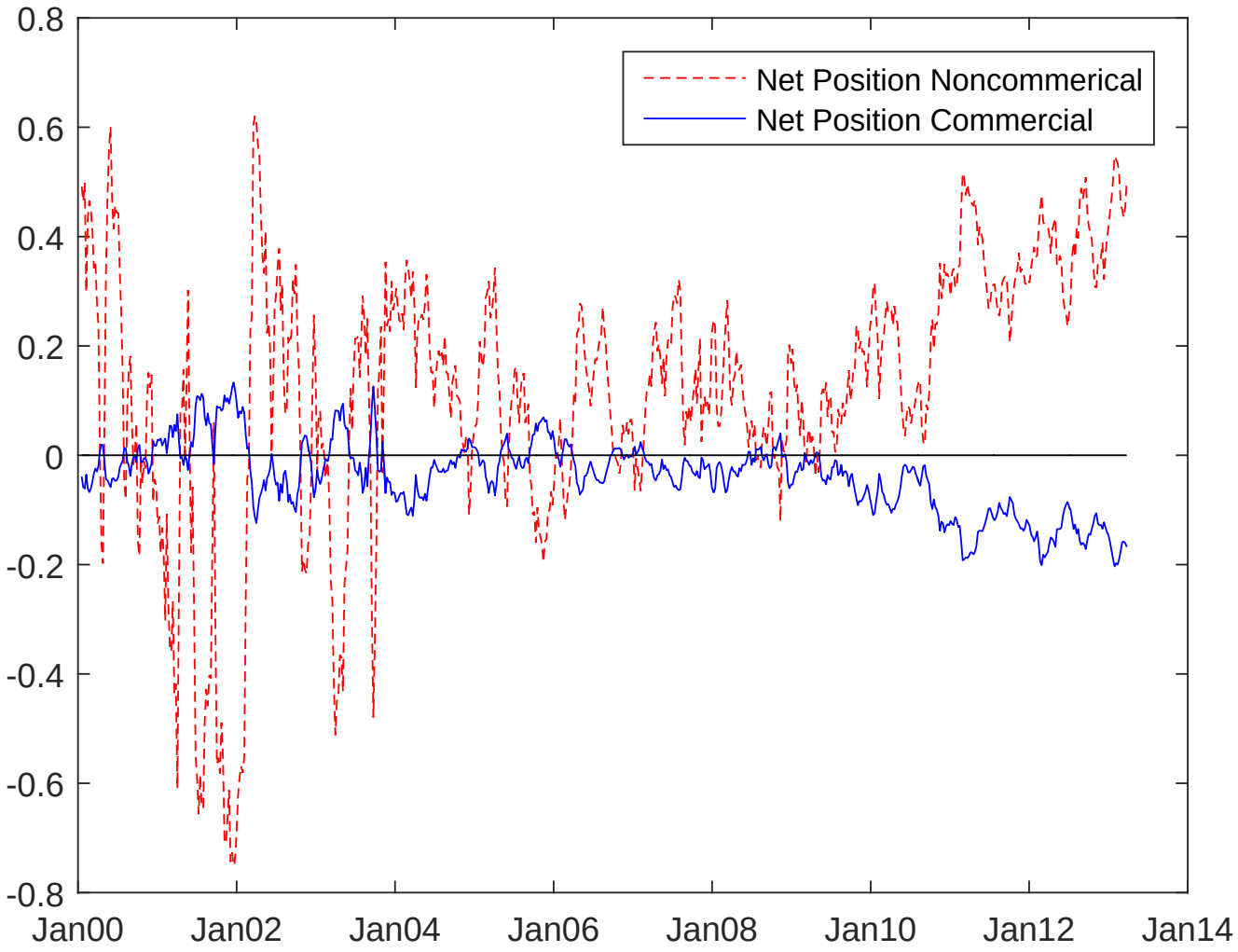

Fig. 6. Data Release and News Sentiment

This figure illustrates the news index of durable goods orders (dashed line) and the actual released durable goods orders index (solid line). Both series are standardized to have zero mean and unit standard deviation.

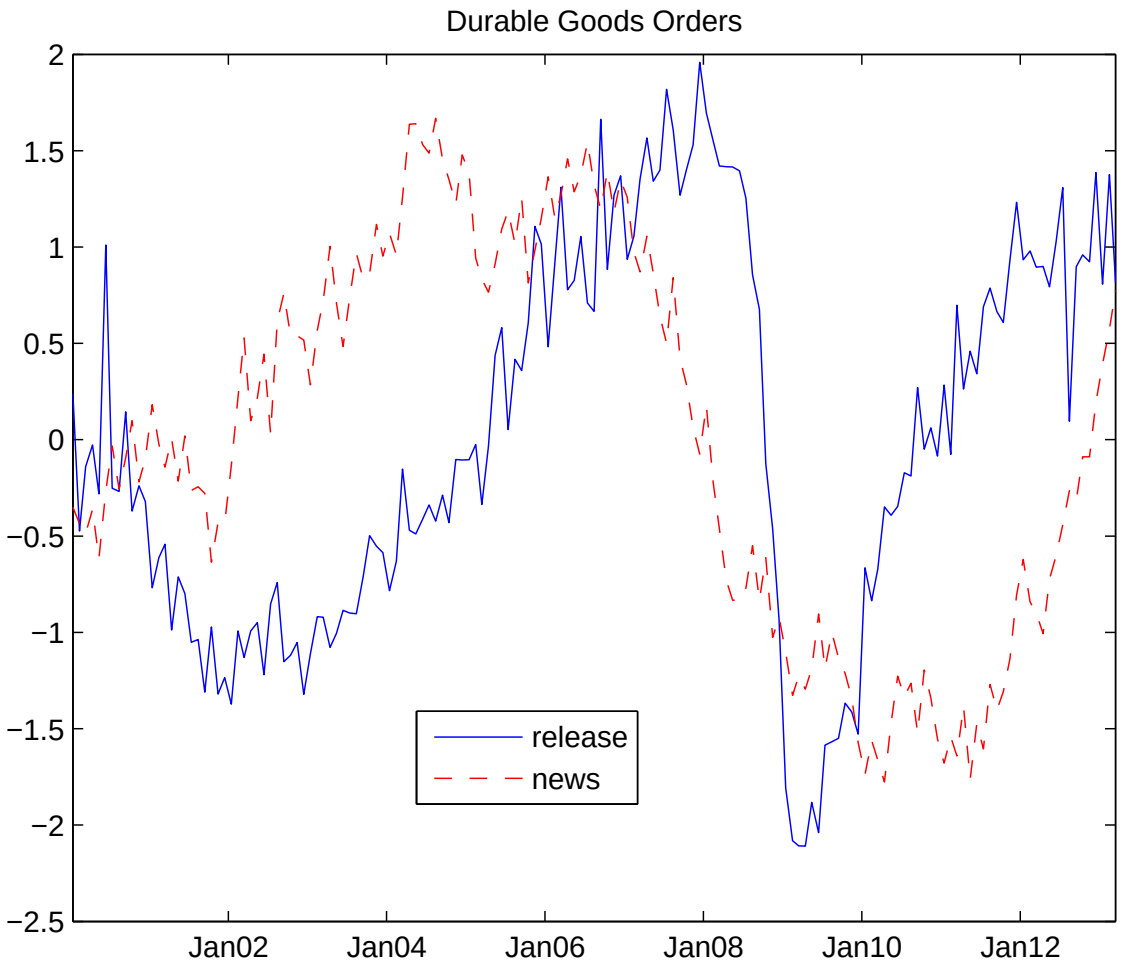

Fig. 7. Data Release and News Sentiment: Crude Oil Suppy

This figure illustrates the news index of crude oil supply (dashed line) and the actual released data of US ending stocks excluding SPR of crude oil (solid line). Both series are corrected for the time trend and are standardized to have zero mean and unit standard deviation.

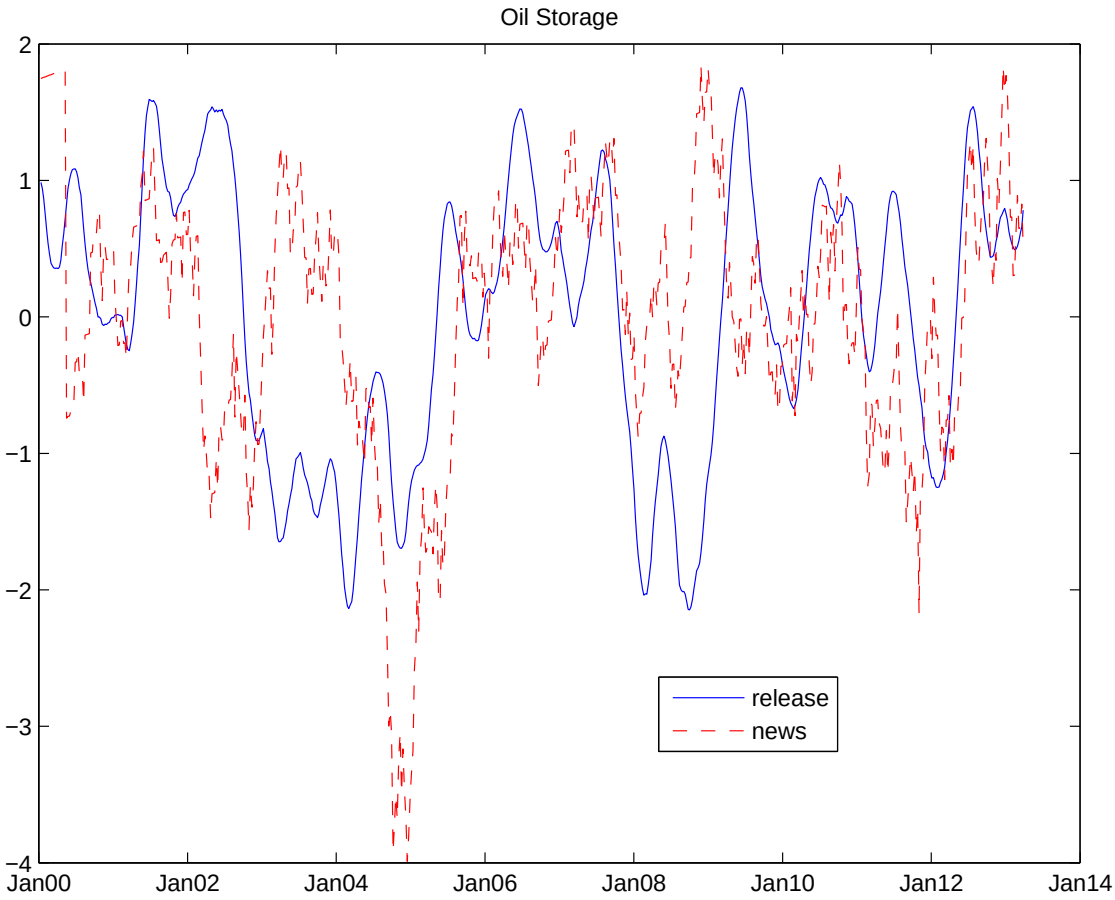

Table 1: Summary Statistics - Crude Oil Returns and News Sentiment Categories

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	335	-12.10	-19	15.54	-40	27
natural disasters	216	-9.63	-13	11.74	-21	25
government	662	-4.11	-16	23.13	-35	35
sovereign debt	161	-3.75	-15	26.85	-41	34
public finance	1051	-6.54	-20	28.42	-50	50
retail sales	1699	6.73	8	26.54	-50	50
credit ratings	1893	3.77	8	21.53	-26	27
consumer confidence	834	1.16	20	32.73	-44	45
foreign exchange	57277	-0.11	-33	33.00	-37	37
housing	516	-0.92	-4	26.33	-43	43
interest rates	561	-3.56	-39	38.90	-47	43
treasury yield	1279	0.53	10	11.12	-32	32
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	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	1128	7.72	31	36.86	-49	50
trade balance	1394	-8.31	-18	28.00	-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	84306	1.30	8	33.32	-50	50
crude oil	3336	0.0004	0.0010	0.0229	-0.1654	0.1334

Table 2: Category Classification

Geopolitical Events:				
terrorism	war & conflict	civil unrest	natural disasters	government
Macro News:				
sovereign debt	public finance	retail sales	consumer confidence	credit ratings
foreign exchange	housing	interest rates	treasury yield	private credit
employment	consumer spending	recession	durable goods orders	economic growth
GDP growth	CPI	PPI	trade balance	exports
Oil Supply and Demand:				
crude oil supply	crude oil demand	price target	drilling & pipeline accident	

Table 3: 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 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	0.0199	0.0537	0.0736	0.0801	0.0185	0.0481	0.0666	0.0721
pct	0.2704	0.7296	1		0.2781	0.7219		
monthly	0.0799	0.2601	0.3399	0.4268	0.0555	0.1594	0.2149	0.2970
pct	0.2349	0.7651	1		0.2583	0.7417		

Table 4: Daily Oil Returns and News

This table reports news coefficients from the regression of daily log oil returns on the news scores in each category, as described in Equation (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 in italics: statistical significance at 10% level; bold: 5% level; italics and bold: 1% level.

	1	2	3	4	5	6	7	8
terrorism	***-12.719***	***-13.095***	***-13.352***	***-12.489***	***-13.569***	***-12.840***	***-13.083***	***-13.264***
war & conflict	***-14.410***	***-13.976***	***-13.842***	***-14.171***	***-13.724***	***-13.671***	***-13.614***	***-13.406***
civil unrest	***-13.002***	***-12.450***	***-12.964***	***-13.082***	***-12.793***	***-12.492***	***-13.189***	***-12.926***
natural disasters	***-12.795***	***-12.859***	***-14.877***	***-13.180***	***-15.148***	***-13.249***	***-15.338***	***-15.673***
government	***21.300***	***21.300***	***17.788***	***20.297***	***17.847***	***20.306***	***17.076***	***17.142***
sovereign debt	***16.403***	***16.465***	***12.703***	***16.110***	***12.934***	***16.185***	***12.614***	***12.912***
public finance	***13.398***	***13.662***	**10.104**	***13.604***	***10.520***	***13.907***	***10.414***	***10.905***
retail sales	***15.254***	***15.247***	***13.402***	***15.759***	***13.336***	***15.759***	***13.646***	***13.603***
credit ratings	***20.008***	***20.034***	***16.912***	***18.886***	***17.329***	***18.937***	***15.993***	***16.429***
consumer confidence	3.786	3.660	2.255	3.058	2.047	2.895	1.521	1.272
foreign exchange	***37.256***	***37.935***	***31.231***	***25.208***	***32.299***	***25.896***	***19.882***	***20.987***
housing	-1.412	-1.346	2.835	-1.641	2.976	-1.591	2.559	2.726
interest rates	**-9.196**	**-8.753**	***-10.691***	**-9.694**	***-10.575***	**-9.195**	***-10.782***	***-10.609***
treasury yield	**-7.573**	*-7.032*	*-6.918*	*-7.441*	*-6.681*	*-6.860*	*-6.647*	-6.329
private credit	***13.322***	***13.132***	***13.929***	***12.823***	***14.014***	***12.596***	***12.372***	***12.465***
employment	***14.522***	***14.045***	**8.770**	***14.470***	**9.035**	***14.013***	**9.033**	**9.278**
consumer spending	2.567	1.904	3.319	1.528	2.867	0.864	2.619	2.119
durable goods orders	-1.568	-1.672	-3.124	-2.146	-3.320	-2.228	-3.421	-3.597
recession	0.828	0.377	-3.148	-0.250	-2.713	-0.646	-4.097	-3.613
economic growth	-1.624	-2.273	-0.401	-0.874	-0.530	-1.532	-0.099	-0.301
GDP growth	5.886	5.522	1.724	5.472	1.977	5.126	1.642	1.881
CPI	3.128	2.350	6.244	4.066	6.017	3.303	6.736	6.516
PPI	-1.787	-2.466	-6.523	-1.783	-5.639	-2.383	-6.670	-5.568
exports	1.874	1.509	-0.321	1.963	-0.226	1.611	-0.588	-0.497
trade balance	-0.821	-1.115	0.320	-0.925	0.419	-1.261	-0.223	-0.068
supply	**-8.808**	**-8.434**	-6.447	**-9.656**	-6.135	**-9.230**	-6.710	-6.380
demand	***11.649***	***11.319***	***11.933***	***11.555***	***12.246***	***11.238***	***11.522***	***11.817***
price target	*7.700*	*7.074*	*7.219*	*7.560*	*7.219*	*6.957*	*6.838*	*6.792*
drilling & pipeline accident	***-13.707***	***-14.322***	***-13.398***	***-13.703***	***-13.763***	***-14.362***	***-13.353***	***-13.771***
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 5: Monthly Oil Returns and News

This table reports the regression of monthly log oil returns on the monthly news scores in each category, as described in Equation (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 in italics: statistical significance at 10% level; bold: 5% level; italics and bold: 1% level.

	1	2	3	4	5	6	7	8
terrorism	0.109	-0.255	0.596	0.109	0.394	-0.255	0.596	0.394
war & conflict	-0.340	-0.009	0.054	-0.340	-0.053	-0.009	0.054	-0.053
civil unrest	*-1.220*	-0.908	-0.889	-1.220	-0.876	-0.908	-0.889	-0.876
natural disasters	***-1.728***	***-1.838***	***-1.795***	***-1.728***	***-1.952***	***-1.838***	***-1.795***	***-1.952***
government	***1.736***	**1.759**	**1.528**	***1.736***	*1.387*	**1.759**	**1.528**	*1.387*
sovereign debt	0.562	0.557	0.315	0.562	0.124	0.557	0.315	0.124
public finance	0.808	1.033	0.556	0.808	0.823	1.033	0.556	0.823
retail sales	0.734	0.828	0.288	0.734	-0.142	0.828	0.288	-0.142
credit ratings	*2.160*	**2.344**	1.002	*2.160*	1.238	**2.344**	1.002	1.238
consumer confidence	0.499	0.362	0.528	0.499	0.243	0.362	0.528	0.243
foreign exchange	***2.724***	***3.182***	**1.773**	***2.724***	***2.494***	***3.182***	**1.773**	***2.494***
housing	-0.439	-0.161	-0.065	-0.439	0.383	-0.161	-0.065	0.383
interest rates	-1.410	-1.117	*-1.457*	*-1.410*	*-1.423*	-1.117	*-1.457*	*-1.423*
treasury yield	***-2.165***	**-1.899**	**-1.901**	***-2.165***	**-1.821**	**-1.899**	**-1.901**	**-1.821**
private credit	0.345	0.251	0.010	0.345	0.194	0.251	0.010	0.194
employment	1.290	0.847	0.378	1.290	0.470	0.847	0.378	0.470
consumer spending	*1.607*	1.245	0.940	*1.607*	0.547	1.245	0.940	0.547
durable goods orders	-0.854	-0.954	-0.460	-0.854	-0.569	-0.954	-0.460	-0.569
recession	*2.562*	**2.859**	0.645	*2.562*	0.893	**2.859**	0.645	0.893
economic growth	*1.770*	**1.824**	1.333	1.770	*1.463*	**1.824**	1.333	*1.463*
GDP growth	1.656	**2.392**	**1.662**	1.656	**2.049**	**2.392**	**1.662**	**2.049**
CPI	**1.889**	**1.529**	**1.456**	**1.889**	1.045	**1.529**	**1.456**	1.045
PPI	-0.292	-0.497	**-1.411**	-0.292	-0.836	-0.497	**-1.411**	-0.836
exports	*1.605*	**1.800**	0.618	*1.605*	0.916	**1.800**	0.618	0.916
trade balance	0.710	0.747	0.096	0.710	0.106	0.747	0.096	0.106
supply	*-1.735*	*-1.524*	**-1.593**	**-1.735**	-1.278	*-1.524*	**-1.593**	-1.278
demand	***3.013***	***3.206***	**1.648**	***3.013***	***2.180***	***3.206***	**1.648**	**2.180**
price target	***4.349***	***4.555***	***2.499***	***4.349***	***2.666***	***4.555***	***2.499***	***2.666***
drilling & pipeline accident	0.959	0.666	0.701	0.959	0.507	0.666	0.701	0.507
control variables+monthly dummies		yes			yes	yes		yes
without crisis period			yes		yes		yes	yes
with SDR				yes		yes	yes	yes

Table 6: Daily Oil Returns and News: Predictive Regressions with Control Variables

This table reports the regression results from Equation (2). Column 1 reports the news categories. Column 2–6 reports the coefficients β_2 in the regression equations. 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.

	$news_{t-1}$	$news_{t-2}$	$news_{t-3}$	$news_{t-4}$	$news_{t-5}$	$adj.R^2$
terrorism	4.277	0.243	-0.077	5.953	4.704	0.0058
war & conflict	**-7.106**	4.633	-0.512	1.917	-0.158	0.0057
civil unrest	-2.322	-3.600	3.988	0.521	5.428	0.0055
natural disasters	-3.549	-2.418	**-7.728**	2.473	-0.304	0.0060
government	-5.773	**7.305**	-1.668	-0.681	***11.843***	0.0087
sovereign debt	-1.276	**7.718**	-4.063	***-10.992***	5.886	0.0086
public finance	-2.622	***9.749***	*-7.808*	3.780	2.246	0.0078
retail sales	6.020	4.721	**8.016**	-3.009	2.628	0.0069
credit ratings	5.025	1.358	**9.228**	-0.067	-3.538	0.0066
consumer confidence	5.579	4.171	-1.690	***11.610***	-3.012	0.0081
foreign exchange	-0.072	3.563	-4.546	-1.180	-1.378	0.0050
housing	-4.012	-5.165	1.052	***12.719***	-2.504	0.0082
interest rates	**-8.142**	0.643	4.590	-0.505	0.478	0.0059
treasury yield	**-8.529**	-4.773	-5.034	0.192	1.560	0.0067
private credit	3.305	0.758	0.888	-0.955	5.037	0.0049
employment	-0.146	4.703	6.178	-1.566	5.548	0.0061
consumer spending	-0.735	-4.106	4.245	-0.444	-6.074	0.0056
durable goods orders	-1.434	3.179	0.017	6.044	0.711	0.0052
recession	**12.844**	-2.921	-2.373	**10.447**	-1.864	0.0101
economic growth	0.649	-1.728	**9.870**	5.458	-4.267	0.0071
GDP growth	2.023	-1.025	-2.188	*6.637*	-1.216	0.0053
CPI	3.585	3.340	-0.686	-0.863	4.278	0.0051
PPI	-3.155	0.781	4.728	-3.060	1.241	0.0051
exports	2.309	4.715	2.112	4.690	-2.397	0.0054
trade balance	-1.259	3.461	3.573	-0.779	2.267	0.0049
supply	2.012	3.121	3.280	-2.506	-5.000	0.0053
demand	-1.138	*7.110*	-0.256	3.226	4.406	0.0059
price target	0.692	-0.179	1.614	3.324	0.936	0.0046
drilling & pipeline accident	0.671	-0.010	**7.913**	-5.351	3.564	0.0062

Table 7: Monthly Oil Returns and News: Predictive Regressions with Control Variables

This table reports the regression results from Equation (2). Column 1 reports the news categories. Column 2-6 reports the coefficients β_2 in the regression equation. Newey-West standard errors are used to calculate the t-statistical significance. Oil returns are expressed in percentage points. News scores are standardized to have zero mean and unit variance. Numbers in italics: statistical significance at 10% level; bold: 5% level; italics and bold: 1% level.

	$news_{t-1}$	$news_{t-2}$	$news_{t-3}$	$news_{t-4}$	$news_{t-5}$	$news_{t-6}$	$adj.R^2$	$\sum_{L=1}^{6}\beta(L)$	Wald(6)	$\sum_{L=1}^{2}\beta(L)$	Wald(2)
terrorism	-0.567	-1.015	-0.026	-1.055	**2.281**	-0.884	0.0998	-1.265	0.398	-1.582	0.571
war & conflict	-0.299	0.512	0.421	-1.364	0.465	**2.030**	0.1174	1.765	0.775	0.213	0.829
civil unrest	0.747	-0.801	***2.188***	0.210	-0.521	0.929	0.1185	2.751	0.897	-0.054	0.771
natural disasters	0.195	-1.063	*1.906*	**-1.880**	0.473	0.175	0.0776	-0.195	0.709	-0.869	0.508
government	1.216	-0.500	1.106	0.209	-0.200	-0.144	0.0708	1.686	0.993	0.715	0.578
sovereign debt	*-1.369*	-0.382	0.990	-1.587	**-1.826**	***3.824***	0.2416	-0.351	0.455	-1.752	0.340
public finance	-0.464	-0.831	**1.784**	**-1.916**	0.208	-0.008	0.0884	-1.227	0.537	-1.294	0.775
retail sales	1.369	-0.240	0.141	*-1.668*	**-2.161**	0.803	0.1183	-1.756	0.517	1.130	0.435
credit ratings	-0.858	0.862	-0.516	-1.050	1.500	0.578	0.0742	0.516	0.841	0.004	0.647
consumer confidence	-1.189	0.618	0.083	0.226	-0.352	1.586	0.0716	0.972	0.992	-0.571	0.687
foreign exchange	-1.074	0.396	0.565	1.331	0.461	-0.163	0.0715	1.516	0.978	-0.678	0.617
housing	-0.823	0.412	-0.342	-0.017	**-1.467**	1.010	0.0747	-1.227	0.976	-0.411	0.719
interest rates	-0.170	0.145	0.155	-0.320	**-1.648**	0.085	0.0701	-1.753	0.974	-0.026	0.970
treasury yield	0.828	**1.817**	-0.306	-1.302	**-2.083**	0.937	0.1261	-0.109	0.181	2.645	0.227
private credit	0.947	0.211	-0.820	***2.539***	-0.118	-1.117	0.1209	1.641	0.370	1.158	0.692
employment	0.872	-0.714	**2.249**	-0.511	-0.217	-0.563	0.0879	1.116	0.579	0.158	0.814
consumer spending	0.522	*-1.174*	0.952	*-1.602*	0.230	0.290	0.0909	-0.783	0.974	-0.653	0.431
durable goods orders	*1.219*	-0.769	-0.663	0.424	-0.766	0.743	0.0873	0.189	0.993	0.450	0.508
recession	*2.908*	-0.825	0.463	-0.718	-0.980	1.393	0.1199	2.242	0.240	*2.083*	0.098
economic growth	**1.828**	-0.416	0.945	1.663	-0.502	-0.583	0.1208	2.935	0.567	1.412	0.238
GDP growth	**2.896**	0.018	-0.580	**-2.857**	**3.376**	-0.071	0.1625	2.782	0.118	2.914	0.140
CPI	***2.978***	***-2.503***	1.186	1.125	-0.054	0.179	0.1877	2.912	0.377	***0.475***	0.001
PPI	0.600	***3.094***	1.926	-0.411	***-2.728***	0.242	0.1474	2.724	0.704	3.694	0.233
exports	**3.128**	-0.131	0.629	0.335	1.201	**-2.480**	0.1669	2.682	0.107	2.997	0.136
trade balance	-0.139	-0.158	-0.038	***3.234***	-0.674	-0.175	0.1500	2.050	0.258	-0.297	0.978
supply	0.164	-0.528	-0.665	0.621	0.786	-0.694	0.0447	-0.317	0.933	-0.365	0.926
demand	-0.510	0.782	0.352	**2.377**	0.968	0.070	0.1411	4.039	0.698	0.273	0.636
price target	-0.465	*1.525*	-1.110	0.726	-1.307	-0.241	0.0640	-0.873	0.895	1.061	0.243
drilling & pipel. accid	-0.285	0.654	-0.486	1.237	-0.732	*1.622*	0.0873	2.009	0.663	0.368	0.767

Table 8: Daily Trading Volume and News

This table reports the regression of daily log difference of oil trading volume on the contemporaneous news sentiment in each category. The original news scores are examined in the first three columns, whereas Columns 4-6 employ the absolute value of the news scores. Newey-West standard errors are used to calculate the t-statistical significance. The data are daily. Numbers in italics: statistical significance at 10% level; bold: 5% level; italics and bold: 1% level.

	original news score			absolute news score		
	$news_t$	t-stat	$adj.R^2$	$news_t$	t-stat	$adj.R^2$
terrorism	-0.909	-1.042	0.0000	0.986	1.050	0.0001
war & conflict	-1.319	-1.208	0.0004	1.602	1.230	0.0005
civil unrest	0.141	0.162	-0.0003	1.368	1.336	0.0003
natural disasters	1.066	1.370	0.0002	-1.223	-1.452	0.0002
government	-0.394	-0.410	-0.0002	0.529	0.470	-0.0002
sovereign debt	-0.173	-0.305	-0.0003	0.011	0.018	-0.0003
public finance	0.648	0.883	-0.0001	-0.658	-0.669	-0.0002
retail sales	-0.650	-0.732	-0.0001	-0.878	-0.729	-0.0001
credit ratings	-1.074	-1.290	0.0002	0.911	0.829	-0.0001
consumer confidence	0.021	0.027	-0.0003	***2.691***	2.689	0.0021
foreign exchange	**-1.876**	-2.401	0.0012	-0.184	-0.119	-0.0003
housing	0.913	0.885	0.0000	1.122	1.598	0.0002
interest rates	0.553	0.629	-0.0002	*1.660*	1.679	0.0007
treasury yield	**1.680**	2.097	0.0009	***3.505***	3.519	0.0035
private credit	-0.714	-1.021	-0.0001	-0.652	-0.876	-0.0001
employment	-0.053	-0.069	-0.0003	0.047	0.044	-0.0003
consumer spending	-0.862	-0.951	0.0000	***-4.538***	-5.055	0.0078
durable goods orders	0.014	0.019	-0.0003	-0.845	-1.127	0.0000
recession	0.030	0.032	-0.0003	-0.003	-0.002	-0.0003
economic growth	-0.932	-1.158	0.0001	*-2.474*	-1.871	0.0006
GDP growth	0.883	0.965	0.0000	1.417	0.846	-0.0001
CPI	***-4.660***	-5.110	0.0088	***-8.389***	-7.098	0.0171
PPI	-0.716	-1.012	-0.0001	-0.539	-0.709	-0.0002
exports	*-1.340*	-1.760	0.0005	0.072	0.074	-0.0003
trade balance	***-2.454***	-2.648	0.0022	-0.296	-0.256	-0.0003
supply	1.014	1.299	0.0001	-1.257	-1.181	0.0000
demand	1.260	1.624	0.0004	***3.007***	3.040	0.0024
price target	-0.514	-0.749	-0.0002	0.107	0.144	-0.0003
drilling & pipeline accident	-0.723	-0.725	-0.0001	0.634	0.608	-0.0001

Table 9: Daily Trading Volume and News: Predictive Regressions

This table reports the VAR estimates in equation (3). The coefficientβ, 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. ΔVlm is expressed in percentage points. The data are daily. Numbers in italics: statistical significance at 10% level; bold: 5% level; italics and bold: 1% level.

	$news_{t-1}$	$news_{t-2}$	$news_{t-3}$	$news_{t-4}$	$news_{t-5}$	$adj.R^2$	$\sum_{L=1}^{5}\beta(L)$	Wald(5)
terrorism	0.524	**1.244**	0.373	0.068	-0.514	0.1741	1.695	0.984
war & conflict	-0.671	*1.698*	0.512	-0.764	0.116	0.1746	0.891	0.955
civil unrest	0.318	*1.115*	0.123	*-1.411*	-0.779	0.1746	-0.634	0.936
natural disasters	0.717	**1.251**	-0.191	*-1.467*	-0.054	0.1746	0.257	0.729
government	*-1.581*	1.220	-0.316	-0.085	0.717	0.1749	-0.045	0.902
sovereign debt	-1.093	**-1.707**	*-1.365*	***-1.960***	-0.316	0.1775	-6.441	0.681
public finance	-0.898	0.399	0.482	0.862	0.534	0.1740	1.379	0.994
retail sales	-0.586	-0.092	0.308	1.007	**1.808**	0.1746	2.445	0.956
credit ratings	0.564	0.323	0.109	0.301	1.005	0.1737	2.303	0.996
consumer confidence	***-2.435***	***-3.221***	***-4.095***	***-3.384***	*-1.511*	0.1845	*-14.645*	0.083
foreign exchange	0.543	**3.537**	-0.337	0.147	-0.969	0.1751	2.921	0.871
housing	-0.546	**-1.763**	**-1.967**	**-3.363**	**-7.403**	**0.2060**	***-15.042***	0.000
interest rates	-0.454	-0.551	0.721	0.712	0.234	0.1738	0.662	0.988
treasury yield	0.272	-0.707	***-4.153***	0.811	***3.381***	0.1839	-0.396	0.512
private credit	***-10.715***	***-5.994***	***-2.467***	***-1.949***	***2.627***	0.2369	***-18.498***	0.000
employment	0.207	-0.416	***3.703***	***5.660***	1.385	0.1855	10.538	0.158
consumer spending	***-6.963***	**-1.959**	**-1.643**	-0.578	0.524	0.1954	-10.619	0.019
durable goods orders	***-2.092***	***-2.465***	***-6.555***	***-4.815***	***-8.624***	0.2359	***-24.551***	0.000
recession	-0.477	-0.726	0.462	*1.384*	0.260	0.1742	0.904	0.961
economic growth	*-2.220*	*-2.366*	*-2.246*	0.461	-1.667	0.1761	-8.037	0.840
GDP growth	-1.997	0.304	-0.156	0.023	1.831	0.1742	0.006	0.975
CPI	***-7.037***	***-4.265***	***-5.207***	***-3.613***	-1.250	0.1975	***-21.372***	0.003
PPI	***-1.599***	0.173	***4.882***	***2.844***	***3.461***	0.1889	**9.761**	0.016
exports	***-3.759***	***-3.046***	*-1.922*	-0.089	-1.331	0.1816	-10.147	0.317
trade balance	-0.984	1.126	0.959	***2.124***	***2.672***	0.1783	5.897	0.583
supply	1.608	-0.382	**-1.996**	1.426	-1.538	0.1755	-0.882	0.873
demand	***2.386***	0.339	0.349	**2.095**	***2.777***	0.1791	7.946	0.320
price target	-0.524	-0.491	-0.530	0.084	0.208	0.1736	-1.254	0.998
drilling & pipeline accident	-0.394	0.755	-1.271	-0.082	-0.154	0.1742	-1.146	0.989

Table 10: Net Noncommercial Traders Positions and News

This table reports the news impact on the net long noncommercial traders positions as in Equation (5). The coefficient β, 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 Equation (4) 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	**0.965**	0.472	0.072	-0.286	0.903	1.223	0.706	1.437	0.291
war & conflict	*0.598*	0.088	0.080	-0.004	0.901	0.763	0.837	0.686	0.412
civil unrest	-0.256	-0.111	0.269	**0.661**	0.902	0.563	0.691	-0.367	0.673
natural disasters	-0.125	-0.004	-0.084	*0.474*	0.901	0.261	0.793	-0.129	0.923
government	0.393	0.049	0.050	*0.560*	0.901	1.051	0.815	0.441	0.573
sovereign debt	0.286	0.127	-0.060	-0.117	0.901	0.236	0.980	0.413	0.718
public finance	0.336	-0.207	-0.201	0.072	0.901	0.000	0.926	0.129	0.549
retail sales	0.386	0.010	0.356	-0.169	0.901	0.583	0.946	0.396	0.693
credit ratings	0.265	-0.021	-0.463	*-0.554*	0.902	-0.772	0.604	0.244	0.785
consumer confidence	0.083	*0.439*	**0.765**	***-0.871***	0.903	0.416	0.271	0.523	0.511
foreign exchange	0.119	0.022	-0.132	***-0.789***	0.902	-0.780	0.711	0.140	0.957
housing	***0.736***	*-0.459*	*0.544*	0.044	0.902	0.865	0.337	0.277	0.027
interest rates	***-0.782***	0.188	0.055	-0.040	0.902	-0.580	0.555	-0.594	0.042
treasury yield	0.018	-0.129	-0.355	0.125	0.901	-0.341	0.969	-0.111	0.942
private credit	-0.358	0.094	**-0.589**	**0.642**	0.902	-0.212	0.802	-0.264	0.779
employment	0.583	0.487	0.361	0.340	0.902	1.771	0.822	1.070	0.373
consumer spending	0.394	-0.085	0.219	-0.199	0.901	0.329	0.930	0.309	0.633
durable goods orders	0.051	**0.713**	0.305	*-0.388*	0.902	0.682	0.697	0.764	0.242
recession	0.386	0.165	0.061	0.009	0.901	0.621	0.952	0.551	0.593
economic growth	0.515	**0.598**	0.040	0.299	0.902	1.451	0.730	1.112	0.249
GDP growth	-0.295	0.368	-0.306	*0.633*	0.901	0.399	0.598	0.073	0.536
CPI	0.484	0.077	-0.371	*-0.523*	0.901	-0.333	0.892	0.561	0.672
PPI	*-0.507*	-0.211	-0.130	-0.120	0.901	-0.969	0.934	-0.719	0.458
exports	0.010	-0.168	0.130	0.084	0.901	0.056	0.999	-0.158	0.947
trade balance	0.309	0.216	-0.306	0.522	0.901	0.741	0.946	0.525	0.783
supply	-0.346	*-0.443*	-0.350	0.337	0.901	-0.803	0.618	-0.789	0.189
demand	-0.129	***0.660***	-0.149	-0.236	0.901	0.146	0.550	0.531	0.077
price target	0.361	-0.243	-0.273	***-0.683***	0.902	-0.837	0.515	0.119	0.486
drilling & pipe. accident	0.394	0.348	-0.094	**-0.550**	0.901	0.097	0.482	0.742	0.332

Table 11: Oil Volatilities and News

This table reports the regression of oil return volatilities on the news sentiment in each category in a similar fashion as in Equation (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 in italics: statistical significance at 10% level; bold: 5% level; italics and bold: 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	**-0.258**	-2.279	0.760	*-0.388*	-1.876	0.329	-0.133	-1.239	0.083
war & conflict	-0.131	-1.222	0.760	*-0.365*	-1.681	0.329	-0.095	-0.902	0.083
civil unrest	***-0.418***	-2.897	0.761	-0.238	-1.086	0.329	**-0.277**	-2.308	0.085
natural disasters	0.056	0.536	0.760	0.221	0.997	0.329	-0.017	-0.184	0.083
government	-0.021	-0.214	0.760	-0.227	-0.967	0.329	0.096	1.212	0.083
sovereign debt	-0.021	-0.204	0.760	-0.140	-0.631	0.329	0.046	0.510	0.083
public finance	**-0.219**	-2.038	0.760	-0.275	-1.155	0.329	***-0.288***	-2.913	0.085
retail sales	**-0.236**	-2.104	0.760	-0.369	-1.521	0.329	**-0.225**	-2.295	0.084
credit ratings	***-0.423***	-3.338	0.761	*-0.431*	-1.786	0.330	***-0.308***	-3.427	0.086
consumer confidence	-0.113	-1.106	0.760	-0.051	-0.243	0.329	-0.088	-0.893	0.083
foreign exchange	***-0.368***	-3.390	0.761	***-0.912***	-3.711	0.332	***-0.240***	-2.644	0.084
housing	0.054	0.455	0.760	-0.268	-1.402	0.329	0.134	1.390	0.083
interest rates	0.072	0.691	0.760	0.255	1.167	0.329	-0.018	-0.187	0.083
treasury yield	0.120	1.089	0.760	0.218	0.853	0.329	0.123	1.497	0.083
private credit	-0.045	-0.492	0.760	0.053	0.304	0.329	-0.043	-0.515	0.083
employment	-0.027	-0.274	0.760	***-0.734***	-3.177	0.331	0.018	0.220	0.083
consumer spending	0.090	0.838	0.760	-0.238	-0.786	0.329	0.011	0.108	0.083
durable goods orders	0.090	0.848	0.760	-0.084	-0.431	0.329	-0.045	-0.471	0.083
recession	**-0.289**	-2.169	0.760	-0.526	-1.579	0.330	-0.132	-1.428	0.083
economic growth	-0.048	-0.478	0.760	0.218	0.904	0.329	-0.053	-0.652	0.083
GDP growth	-0.091	-0.750	0.760	0.052	0.204	0.329	-0.065	-0.594	0.083
CPI	-0.049	-0.492	0.760	0.060	0.262	0.329	0.008	0.087	0.083
PPI	0.001	0.013	0.760	-0.120	-0.461	0.329	0.003	0.034	0.083
exports	-0.067	-0.677	0.760	0.289	1.234	0.329	-0.104	-1.216	0.083
trade balance	-0.075	-0.681	0.760	0.070	0.315	0.329	-0.017	-0.186	0.083
supply	0.051	0.534	0.760	**0.450**	2.122	0.330	-0.038	-0.467	0.083
demand	**-0.241**	-2.088	0.760	**-0.451**	-2.273	0.330	-0.009	-0.095	0.083
price target	-0.072	-0.550	0.760	-0.780	-2.709	0.331	***0.303***	3.418	0.085
drilling & pipe. accident	-0.053	-0.570	0.760	0.277	1.380	0.329	-0.001	-0.010	0.083

Appendix

Table A.1: Daily Predictive Regressions: Controlling for Crisis period

This table reports the results from the regression equation (1) after excluding the crisis period. Column 1 reports the news categories. Column 2-6 reports the coefficients β_2 in the regression equation. Newey-West standard errors are used to calculate the t-statistical significance. Italics denotes significance at 10% level; bold denotes significance at the 5% level; italics and bold denotes significance at the 1% level.

	$news_{t-1}$	$news_{t-2}$	$news_{t-3}$	$news_{t-4}$	$news_{t-5}$	$adj.R^2$
terrorism	3.202	-4.460	4.458	*9.602*	3.204	0.0013
war & conflict	-6.010	2.049	-0.010	2.057	-0.216	-0.0016
civil unrest	-2.260	-4.981	2.554	-1.787	3.697	-0.0013
natural disasters	0.134	-1.101	**-7.393**	-0.850	-1.049	-0.0009
government	-2.800	5.281	-0.784	0.603	**8.651**	-0.0001
sovereign debt	-1.628	4.738	-4.568	**-8.280**	3.840	0.0002
public finance	**-7.923**	**8.715**	-4.284	**8.381**	1.451	0.0027
retail sales	2.692	*7.309*	**8.594**	0.598	3.151	0.0008
credit ratings	3.147	1.543	3.721	2.279	-3.798	-0.0015
consumer confidence	3.657	3.525	-3.113	***12.135***	-0.190	0.0015
foreign exchange	-3.822	3.410	**-8.230**	-0.286	-4.660	0.0002
housing	-2.353	-5.534	0.746	***14.395***	0.375	0.0027
interest rates	**-7.614**	2.091	1.760	-2.968	3.334	-0.0008
treasury yield	**-9.023**	*-6.140*	*-5.689*	3.049	0.569	0.0013
private credit	5.151	-1.041	0.122	-5.717	1.393	-0.0016
employment	-0.338	1.846	4.099	2.360	6.076	-0.0011
consumer spending	-0.097	-0.148	2.952	-1.899	-4.478	-0.0018
durable goods orders	2.626	-0.486	1.454	4.718	-1.476	-0.0019
recession	5.170	-2.289	0.278	2.959	3.146	-0.0016
economic growth	1.616	-0.210	***10.454***	5.367	-0.163	0.0005
GDP growth	1.248	*-6.931*	-0.293	1.025	-2.160	-0.0014
CPI	2.182	1.302	-5.791	-1.188	**9.826**	0.0004
PPI	-3.679	-4.476	2.235	3.674	-2.079	-0.0013
exports	2.967	1.757	-2.784	0.130	-0.103	-0.0021
trade balance	-1.544	3.192	0.466	-2.565	2.407	-0.0020
supply	0.906	-0.351	-1.152	-0.952	-3.954	-0.0021
demand	-2.096	5.166	0.444	*8.097*	4.070	-0.0001
price target	0.729	2.203	-3.166	1.392	0.559	-0.0022
drilling & pipeline accident	1.384	-1.339	*7.416*	-2.694	*5.040*	-0.0003

Table A.2: Monthly Predictive Regressions: Controlling for Crisis Period and with Control Variables

This table reports the results from the regression equation (2) 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	-0.002	-0.192	-0.715	**-2.257**	*2.007*	-1.215	0.0497	-2.374	0.735	-0.194	0.988
war & conflict	0.050	0.317	0.852	-0.055	0.628	*1.629*	-0.0074	3.420	0.983	0.367	0.966
civil unrest	0.523	*-1.309*	*1.659*	0.602	0.477	0.787	0.0131	2.739	0.951	-0.786	0.547
natural disasters	0.436	-0.520	*1.799*	-1.114	-0.073	0.400	-0.0397	0.927	0.908	-0.084	0.760
government	0.965	0.827	0.461	0.415	-1.134	-0.552	-0.0197	0.983	0.998	1.792	0.765
sovereign debt	-1.234	0.023	**1.944**	-0.409	**-2.049**	***3.330***	0.1182	1.604	0.267	-1.211	0.489
public finance	-0.745	-1.543	*1.600*	**-2.419**	0.791	-0.031	0.0480	-2.346	0.563	-2.288	0.347
retail sales	0.607	-0.661	0.538	-0.598	-0.821	-0.049	-0.0591	-0.983	0.996	-0.054	0.775
credit ratings	**-1.787**	1.587	-0.466	0.021	0.593	0.362	-0.0199	0.310	0.911	-0.200	0.134
consumer confidence	0.456	0.926	0.656	-0.039	-1.448	1.610	-0.0005	2.161	0.982	1.382	0.807
foreign exchange	-0.957	0.555	-0.669	1.000	-0.420	-0.573	-0.0443	-1.065	0.989	-0.402	0.663
housing	0.610	0.738	-0.949	-0.408	*-1.482*	-0.014	-0.0395	-1.505	0.997	1.348	0.803
interest rates	-1.185	1.192	-0.093	0.017	-0.873	-1.351	-0.0114	-2.292	0.863	0.008	0.422
treasury yield	0.467	1.438	-0.727	-0.523	-0.550	-0.063	-0.0496	0.042	0.944	1.904	0.605
private credit	1.290	1.043	0.181	1.861	0.221	-1.746	0.0078	2.851	0.987	2.334	0.617
employment	-0.420	-0.500	**2.396**	0.425	0.651	0.037	0.0016	2.589	0.773	-0.920	0.927
consumer spending	**1.609**	*-1.613*	1.130	**-1.895**	-0.167	-0.623	0.0567	-1.560	0.964	-0.004	0.258
durable goods orders	***2.222***	0.469	1.022	*1.527*	-0.857	0.580	0.0647	4.963	0.944	2.691	0.117
recession	0.326	1.342	0.540	2.029	-2.061	-1.285	-0.0128	0.891	0.089	1.668	0.805
economic growth	1.460	-0.938	-0.166	1.293	0.655	0.441	0.0042	2.745	0.951	0.522	0.295
GDP growth	***4.359***	-1.120	-0.087	**-2.863**	**3.611**	-0.684	0.1510	*3.216*	0.074	***3.239***	0.003
CPI	**2.373**	**-2.592**	1.085	0.516	-0.176	-1.134	0.0840	0.072	0.890	**-0.219**	0.084
PPI	-0.200	***3.289***	**2.730**	0.458	-1.794	-1.305	0.0797	3.177	0.767	3.089	0.255
exports	1.193	0.937	-0.480	1.445	**2.243**	***-3.670***	0.1387	**1.668**	0.011	2.130	0.560
trade balance	0.339	-0.759	-0.559	***2.999***	-0.595	0.717	0.0760	2.142	0.683	-0.421	0.790
supply	0.929	-0.308	-1.369	-0.271	1.397	-0.997	-0.0363	-0.619	0.749	0.621	0.679
demand	-0.469	0.652	1.234	0.648	*2.073*	-1.474	0.0189	2.664	0.668	0.183	0.868
price target	-0.694	-0.633	-0.905	0.946	-1.146	-0.938	-0.0459	-3.371	0.985	-1.327	0.861
drilling & pipe. accid.	-0.186	0.530	-0.794	***2.218***	-1.686	*1.378*	0.0327	1.460	0.141	0.344	0.874

Table A.3: Net Commercial Traders Positions and News

This table reports the news impact on the net long commercial traders positions as in Equation (5). The coefficient β, 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 Equation (4) 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	***-0.251***	-0.078	0.008	0.068	0.941	-0.252	0.340	-0.329	0.079
war & conflict	*-0.106*	-0.065	-0.013	-0.010	0.940	-0.194	0.834	-0.171	0.351
civil unrest	0.075	0.002	-0.005	**-0.168**	0.940	-0.095	0.651	0.078	0.595
natural disasters	0.081	-0.048	0.041	**-0.141**	0.940	-0.066	0.313	0.033	0.369
government	-0.028	-0.041	0.014	-0.087	0.940	-0.142	0.940	-0.069	0.824
sovereign debt	-0.034	-0.015	0.019	-0.019	0.940	-0.049	0.997	-0.049	0.904
public finance	0.005	-0.001	0.027	-0.007	0.940	0.024	0.999	0.004	0.998
retail sales	*-0.109*	-0.017	-0.066	0.060	0.940	-0.133	0.899	-0.126	0.464
credit ratings	-0.065	0.025	*0.099*	**0.128**	0.940	0.187	0.589	-0.040	0.697
consumer confidence	-0.080	-0.046	*-0.129*	**0.136**	0.941	-0.119	0.575	-0.126	0.552
foreign exchange	-0.028	-0.038	0.028	**0.150**	0.940	0.112	0.809	-0.066	0.862
housing	-0.083	0.045	-0.054	-0.081	0.940	-0.173	0.788	-0.038	0.424
interest rates	**0.119**	-0.066	-0.009	-0.025	0.940	0.020	0.681	0.054	0.097
treasury yield	0.010	0.035	0.017	-0.072	0.940	-0.009	0.954	0.045	0.909
private credit	0.118	-0.043	0.102	-0.037	0.940	0.140	0.895	0.075	0.529
employment	-0.121	**-0.135**	-0.102	-0.056	0.941	-0.414	0.692	-0.256	0.283
consumer spending	*-0.096*	0.038	-0.077	0.041	0.940	-0.095	0.885	-0.059	0.533
durable goods orders	-0.007	**-0.138**	-0.054	0.013	0.940	-0.186	0.829	-0.145	0.282
recession	*-0.096*	-0.001	0.014	0.017	0.940	-0.066	0.941	-0.097	0.541
economic growth	-0.072	**-0.146**	0.002	-0.098	0.941	-0.314	0.584	-0.218	0.203
GDP growth	0.102	-0.054	0.052	-0.087	0.940	0.013	0.689	0.048	0.434
CPI	*-0.131*	0.019	0.095	0.031	0.940	0.013	0.834	-0.112	0.390
PPI	**0.113**	0.133	0.016	-0.009	0.940	0.253	0.835	0.245	0.204
exports	0.040	0.038	-0.037	-0.020	0.940	0.020	0.994	0.078	0.843
trade balance	0.006	-0.021	0.054	-0.094	0.940	-0.056	0.961	-0.015	0.974
supply	0.052	0.089	0.055	-0.062	0.940	0.135	0.816	0.142	0.390
demand	-0.062	**-0.131**	0.036	0.039	0.940	-0.118	0.664	-0.194	0.113
price target	-0.065	0.073	0.053	***0.173***	0.941	0.234	0.537	0.008	0.528
drilling & pipe. accident	-0.093	-0.056	0.058	***0.161***	0.941	0.070	0.145	-0.149	0.282

References

Antweiler, W. and Frank, M. Z. (2004). Is all that talk just noise? The information content of internet stock message boards. *The Journal of Finance*, 59(3):1259–1293.

Bakshi, G., Kapadia, N., and Madan, D. (2003). Stock return charakteristics, skew laws, and the differential pricing of individual equity options. *The Review of Financial Studies*, 16(1):101–143.

Bollerslev, T., Tauchen, G., and Zhou, H. (2009). Expected stock returns and variance risk premia. *Review of Financial Studies*, 22(3):4463–4492.

Boudoukh, J., Feldman, R., Kogan, S., and Richardson, M. (2013). Which news moves stock prices? A textual analysis. *NBER Working Paper*.

Breunig, R. V. and 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. *Journal of Financial Economics*, 70(2):223–260.

Cochrane, J. H. and Piazzesi, M. (2005). Bond risk premia. *American Economic Review*, 95(1):138–160.

Coval, J. D. and Shumway, T. (2001). Is sound just noise? *The Journal of Finance*, 56(5):1887–1910.

de Roon, F. A., Nijman, T. E., and Veld, C. (2000). Hedging pressure effects in futures markets. *The Journal of Finance*, 55(3):1437–1456.

Dougal, C., Engelberg, J., Garcia, D., and Parsons, C. (2012). Journalists and the stock markets. *The Review of Financial Studies*, 25(3):639–679.

Garcia, D. (2013). Sentiment during recessions. *The Journal of Finance*, 68(3):1267–1300.

Gomes, J. F., Kogan, L., and Yogo, M. (2009). Durability of output and expected stock returns. *Journal of Political Economy*, 117(5):941–986.

Hamilton, J. (2008). Understanding crude oil prices. *NBER Working Paper*.

Hong, H. and Stein, J. (1999). A unified theory of underreaction, momentum trading and overreaction in asset markets. *The Journal of 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. *American Economic Review*, 99(3):1053–1069.

Kilian, L. and Baumeister, C. (2014). A general approach to recovering market expectations from futures prices with an application to crude oil. *Working Paper*.

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

Peress, J. (2014). The media and the diffusion of information in financial markets: Evidence from newspaper strikes. *The Journal of Finance*, 69(5):2007–2043.

Ready, R. (2013). Commodity trade and the carry trade: A tale of two countries. *NBER Working Paper*.

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. *The Journal of Finance*, 62(3):1139–1168.

Tetlock, P. C. (2011). All the news that's fit to reprint: Do investors react to stale information. *The Review of Financial Studies*, 24(5):1481–1512.

Zhang, S. S. (2013). Need for speed: An empirical analysis of hard and soft information in a high frequency world. *Working Paper*.