Finance and Economics Discussion Series

Federal Reserve Board, Washington, D.C.
ISSN 1936-2854 (Print)
ISSN 2767-3898 (Online)

Reasons Behind Words: OPEC Narratives and the Oil Market

Celso Brunetti, Marc Joëts, Valérie Mignon

2024-003

Please cite this paper as:
Brunetti, Celso, Marc Joëts, and Valérie Mignon (2024). "Reasons Behind Words: OPEC Narratives and the Oil Market," Finance and Economics Discussion Series 2024-003. Washington: Board of Governors of the Federal Reserve System, https://doi.org/10.17016/FEDS.2024.003.

Reasons Behind Words: OPEC Narratives and the Oil Market

Celso Brunetti* Marc Joëts† Valérie Mignon‡

January 10, 2024

Abstract

We analyze the content of the Organization of the Petroleum Exporting Countries (OPEC) communications and whether it provides information to the crude oil market. To this end, we derive an empirical strategy which allows us to measure OPEC's public signal and test whether market participants find it credible. Using Structural Topic Models, we analyze OPEC narratives and identify several topics related to fundamental factors, such as demand, supply, and speculative activity in the crude oil market. Importantly, we find that OPEC communication reduces oil price volatility and prompts market participants to rebalance their positions. Our analysis indicates that market participants assess OPEC communications as providing an important signal to the crude oil market.

JEL Classification: G10, Q35, Q40, C45, C50

Keywords: OPEC Announcements, Structural Topic Models, Volatility, Traders' Positions

Acknowledgements: We would like to thank Christiane Baumeister, Anna Creti, Thomas Grjebine, Lutz Kilian, Julia Silbert, and participants at the CEMA annual meeting, and the 16th CFE conferences for very constructive comments and suggestions.

*Division of Research and Statistics, Board of Governors of the Federal Reserve System, United States of America. E-mail: celso.brunetti@frb.gov

†IESEG School of Management; Univ. Lille, CNRS UMR 9221 - LEM Lille Economie Management F-59000 Lille, France. E-mail: m.joets@ieseg.fr

‡EconomiX-CNRS, University of Paris Nanterre and CEPII, Paris, France. E-mail: valerie.mignon@parisnanterre.fr

1 Introduction

Communication is essential in policy institutions, governmental and intergovermental organizations as well as firms. As the leading entity in oil markets, the Organization of the Petroleum Exporting Countries (OPEC) regularly shares information by releasing communications. OPEC's objective is "[...] to coordinate and unify the petroleum policies of its Member Countries and ensure the stabilization of oil markets [...]".[1] A well-functioning crude oil market may have positive implications for the economy and inflation.[2]

Falling into this context, this paper aims to extract the information content in OPEC communication using textual analysis. In particular, we are interested in identifying the topics of OPEC's announcements, the fundamental factors driving these topics, and how these topics are connected to form OPEC's overall narrative. We are then interested in understanding whether OPEC's narrative percolates into the crude oil market. Traditional finance theory suggests that information flows into markets through volatility and volume (Epps & Epps (1976), and Gallant et al. (1992)). Hence, our objective in the second part of the paper is to understand whether OPEC communication is credible and, if so, how it relates to both oil price volatility and the trading behavior of commercial and non-commercial market participants.

We begin by analyzing all OPEC communications, then use Structural Topic Models (STM)[3] to extract the information content (signal) in OPEC narratives. In line with Morris & Shin (2002), we assume that for OPEC's public signal to be credible, it should reflect market fundamentals. Therefore, to estimate the structural dynamics of OPEC communication, we consider several exogenous factors, such as oil demand, oil supply, and speculative activity in the oil market.[4] We are able to identify numerous topics embedded in OPEC communication, which characterize the Organization's public signal. Besides the obvious topics, such as "prices," "oil shortage," and "economic growth," we detect topics related to "climate change" and "energy policies." This is not surprising, since climate change and climate-related risks have both direct (i.e., new policies to reduce fossil fuels emissions) and indirect (i.e., new technologies) effects on oil-producing countries.

The richness of our textual analysis results allows us to study the rationale behind OPEC communication. We do so in two ways. First, we map the network of topics in OPEC communication in order to investigate the interconnectedness of OPEC announcements and illustrate the complexity of OPEC's narrative. Second, we identify factors that impact OPEC topics. The results of this analysis allow us to better understand the structure of OPEC communi-

[1]See https://www.opec.org

[2]OPEC's production agreements are not binding, and there is no enforcement mechanism. In fact, compliance of OPEC's countries to production agreements has fluctuated over time.

[3]See Roberts et al. (2013).

[4]As it is well-established in the literature, these factors are exogenous with respect to OPEC communications and crude oil price—see, e.g., Kilian (2023) and the references therein.

cation, which is a first step in our quest to test whether OPEC communication contains a credible signal. Our textual analysis shows that OPEC topics are linked to important factors and are clustered in a meaningful manner, suggesting that OPEC narratives are based on crude oil fundamentals.

To test how OPEC narratives impact the oil market, we adopt Lasso penalized regressions. Based on the Morris & Shin (2002) and Amato et al. (2002) theoretical framework, we develop an empirical strategy to crucially test the hypotheses that OPEC's public signal (i) matters to the oil market and (ii) changes as a function of the precision of the private signal. We are aware that endogeneity concerns may affect our analysis. However, our identification strategy is motivated by two observations. First, OPEC is the dominant player in the crude oil market, and its announcements are closely monitored by market participants. Second, our analysis concentrates only on the days and weeks when there are announcements. Therefore, and similar to Känzig (2021), we isolate the impact of OPEC communication and consider only the window around OPEC's announcements.

Our results unequivocally show that OPEC signal is highly related to crude oil market volatility and traders' positions. We find that OPEC communication is associated with lower volatility levels.[5] These results are stronger for longer-dated futures contracts, indicating that OPEC narratives are linked to the entire futures curve. OPEC communication that reassures the market on production capacity and supply contributes to oil market stabilization. Turning to traders' positions, our findings indicate that market participants' trading activity is deeply linked to OPEC topics.[6] In particular, different topics are associated with different traders. Traders engaged in the physical business (i.e., producers, merchants, processors and users) change their positions with OPEC topics related to economic growth, while non-commercial traders (i.e., traders with no business in the underlying physical market), change their positions with topics related to oil supply and "energy policy." All market participants rebalance their positions with the topic "cooperation." Since OPEC does not have an enforcing mechanism, cooperation among members is essential for a credible signal.[7] We also find substantial evidence that OPEC's public signal is stronger when the private signal is noisier, in line with the predictions in Morris & Shin (2002).

We run several robustness checks which confirm our results. Importantly, we also implement a placebo test, consisting of selecting days and weeks randomly with no OPEC announcements and constructing different control groups. We show that, in the absence of OPEC announcements, there is no effect on volatility and traders' positions, providing further support to our

[5] We measure volatility with the daily range, i.e., the difference between the daily log-high and log-low prices.

[6] We use data from the Disaggregated Commitment of Traders (DCOT) reports from the Commodity Futures Trading Commission (CFTC) to measure trading activity of crude oil market participants.

[7] The absence of an enforcing mechanism could be particularly important during crisis periods. In the last part of the paper, we analyze the Global Financial Crisis (GFC) and the COVID-19 pandemic. Our results are robust to the two crises.

identification strategy and confirming the existence of a causal relationship.

The paper provides several contributions. First, we are the first to apply Structural Topic Models to analyze OPEC announcements. This technique allows us to identify relevant topics and study how those topics are connected in a network. Furthermore, we analyze the drivers (supply, demand, and speculative factors) of the estimated topics. Pescatori & Nazer (2022) study OPEC communication and use cosine similarity and term-frequency-inverse document frequency techniques to assess the distance, in terms of content, between different OPEC announcements. Repetitive communications are considered to be uninformative. They find that OPEC statements do not vary much except when oil prices fluctuate dramatically, such as in 2008 when they reached about $150 per barrel. Our approach is different, since we are able to precisely estimate the topics of OPEC narratives and, thus, the information content of OPEC announcements.

Second, to the best of our knowledge, we are the first to apply the Morris & Shin (2002) and Amato et al. (2002) theoretical framework to OPEC communication and to the oil market. A large literature applies this approach to central bank communication,[8] but not to OPEC announcements. Starting from the seminal work of Morris & Shin (2002), we derive testable hypotheses and build an empirical strategy to test them.

Third, we test whether topics in OPEC communication are linked to volatility and traders' positions. We are particularly interested in these two variables because there is substantial evidence showing how trading volume and volatility are related through the information flow.[9] We consider crude oil price volatility over the entire maturity futures curve and disaggregated traders' positions from the CFTC public data. There is a large literature studying the credibility of OPEC communication with mixed results. Wirl & Kujundzic (2004) find weak evidence of the impact of OPEC communication on the world oil market, while Guidi et al. (2006) show that the effectiveness of OPEC decisions varies over time. Similarly, Demirer & Kutan (2010), using an event study approach, find that only OPEC production cut announcements have an impact on oil prices, and this impact vanishes for longer maturity futures contracts. In a similar camp are the results of Fattouh & Mahadeva (2013), which show that OPEC pricing power varies over time. Brunetti et al. (2013) provide empirical evidence that OPEC fair price pronouncements have limited effects on the actual price of crude oil. Looking at more recent data, Quint & Venditti (2020) also find a limited effect of OPEC+ on crude oil markets. Some studies, however, provide substantive evidence that OPEC announcements have a significant influence on crude oil markets. Lin & Tamvakis (2010) and Loutia et al. (2016), using an event study approach, find significant crude oil market responses to OPEC production decisions.[10]

[8] See, among others, Hermann & Fratzscher (2007a), Hermann & Fratzscher (2007b), and Evans et al. (2012).

[9] See, Epps & Epps (1976), Tauchen & Pitts (1983), and Gallant et al. (1992).

[10] Loutia et al. (2016) account for the volatility structure of crude oil prices using an Exponential GARCH model.

Our approach is different, as we do not use an event study methodology. Rather, we precisely measure the topics in OPEC communication. Our results suggest that OPEC communication is based on fundamental factors and generates a credible public signal. In particular, we find that OPEC topics reduce volatility levels, in line with OPEC's mandate of market stabilization, and induce market participants to rebalance their positions. This is particularly true when the private signal is noisy, in accordance with the predictions in Morris & Shin (2002).

The remainder of the paper is organized as follows. Section 2 is devoted to the information content of OPEC narratives. Section 3 presents the theoretical framework and the related empirical strategy to test the effectiveness of OPEC communication. Section 4 studies the linkages between OPEC topics and volatility and traders' positions. Section 5 provides several robustness checks, including a sensitivity analysis to recent crisis periods. Section 6 concludes the paper.

2 The information content of OPEC communication

This section begins by introducing the database of OPEC press releases, as well as identifying the fundamental drivers influencing OPEC communications. Next, we outline the methodology employed to quantify the information content in OPEC statements through textual analysis. Finally, we examine the temporal evolution of OPEC narratives, explore their interconnections, and identify the driving forces behind each topic.

2.1 OPEC press releases and fundamental drivers

The information content of OPEC communication is estimated using textual analysis, as described in Section 2.2. Our base corpus consists of OPEC press releases extracted from OPEC's website, starting from March 2002 and updated each time a representative member gives an official talk to the press.[11] Declarations from the Organization to the press usually come around major events, such as ordinary and extraordinary OPEC conferences, but also during episodes of oil market turbulence. In this respect, our corpus is quite imbalanced, with years of intense communications (more than one per month) followed by periods of limited communications.

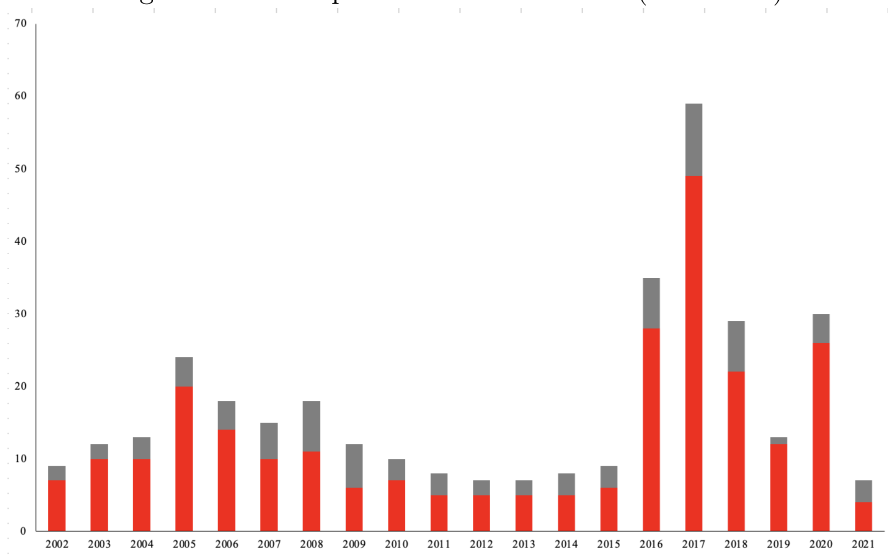
Considering every press release from March 2002 to March 2021, the sample includes 343 announcements.[12] Figure 1 reports in grey the number of announcements per year over the sample period and shows how OPEC communication is cyclical. As expected, periods of intense communication coincide with episodes of prolonged low oil prices, such as the development of the US shale oil in 2016-2017 and the COVID-19 pandemic in 2020. From the total speeches, we drop talks with no topical content (such as data exercises or administrative

[11] See www.opec.org/opec_web/press_room/28.htm.

[12] The detailed database including dates, speakers, locations, and titles of each press release, is available upon request from the authors.

issues), leaving 262 press releases over our sample period (in red in Figure 1).

Figure 1: OPEC press releases over time (2002-2021)



Note: This figure depicts the number of annual OPEC press releases from March 2002 to March 2021 (in grey), totaling 343 speeches (source: OPEC's website), with our selected sample of 262 speeches highlighted in red.

As highlighted in the existing literature, OPEC continuously assesses the oil market to establish targeted prices and appropriate supply levels (Kilian & Murphy (2014), Baumeister & Kilian (2016)). Consequently, we posit that OPEC's signaling is endogenous to market conditions, with the Organization observing fundamental factors in the crude oil market prior to issuing a statement. In order to estimate the endogenous structural dynamics influencing OPEC's communications, we incorporate several exogenous variables that could impact the Organization's messaging to the market. These variables enable us to consider multiple market components—including demand, supply, and speculative factors—and serve to identify the topics contained in OPEC's public statements.[13] These variables were constructed for the period from March 2002 to March 2021 to align with the frequency of OPEC press releases and are elaborated upon below.

While some of the considered exogenous factors are quantitative in nature (i.e., evolve over time), we transform all variables to be qualitative, categorizing them into groups or classes. The transformation is performed for two reasons: (i) analytically, we are interested in OPEC signal with respect to some reference point, for instance when the US business conditions improve (or worsen), and (ii) technically, computation and interpretation are made easier when working with qualitative variables.

[13] See, e.g., Kilian & Murphy (2014), Baumeister & Kilian (2016), and Brunetti et al. (2016) for a discussion.

We consider a total of eight exogenous qualitative variables organized into three categories: supply, demand, and speculative. Regarding supply, we rely on spare crude oil capacity. We construct a two-class variable, which we label "low" when capacity is low or close to zero (for instance between 2004-2008 and between 2012-2018) and "high" otherwise. We build five demand-related variables:

- US business conditions: we use the ADS index developed by Aruoba et al. (2009) to proxy US business conditions. By construction, the average value of the index is zero, and progressively larger positive (negative) values indicate progressively better-than-average (worse-than-average) conditions. We construct a two-class variable capturing "worse" conditions if negative, and "better" if positive.

- Economic uncertainty for the US and Europe: we use the uncertainty index of the state of the economy developed by Scotti (2016). We construct a two-class variable which we label "high" when the index is higher than 1.65 standard deviation, and "low" otherwise.

- Economic surprise for the US and Europe: we use the surprise index developed by Scotti (2016) summarizing recent economic data surprises and measures of optimism and pessimism about the state of the economy. We construct a two-class variable recording "optimism" when negative, and "pessimism" when positive.

Finally, we construct two variables that capture speculative activity, namely US & OECD crude oil inventories. We use US crude oil stocks from Energy Information Administration and OECD crude oil stocks from the International Energy Agency. We then construct a two-class variable for each inventory considering "high" when stocks are outside the 5-year average band (for instance during the shale oil period of 2015-2017), and "normal" otherwise. The use of inventories to capture speculative activity has been considered in several studies, such as Kilian & Murphy (2014).

2.2 Structural Topic Model

For topics estimation, we rely on unsupervised probabilistic topic models applied to OPEC press releases. Among these models, mixed-membership approaches allowing each document to be composed of multiple topics, have become a common tool for mining large corpora in various fields.[14] The intuition, popularized by Blei et al. (2003)'s Latent Dirichlet Allocation (LDA), is that a document is a collection of multiple topics, which are themselves a collection of words. A topic is then defined as a mixture of words where each word has a probability of belonging to a topic. A document is a mixture of topics, and so a single document can be composed of multiple topics.

The main goal of these models is to estimate the following three components:

- Topic proportions (i.e., document-topic probability distributions) for each document $d \in \{1, 2..., D\}$ (also called topic prevalence) described by the parameter θ_d.

[14] See Blei & Lafferty (2009) and Blei (2012) for a review.

- Word proportions (i.e., topic-word probability distributions) for each topic $k \in \{1, 2..., K\}$ (also called topic content) described by the parameter β_k.
- Core language combining the two previous components to produce the actual words in each document. In other terms, for each word $n \in \{1, 2..., N\}$ in document d, a per-word topic assignment z conditional on the document-topic probability distribution is drawn from a multinomial distribution $(z_{d,n} \, |\theta_d)$. Given the topic, words are randomly chosen from a multinomial distribution $(w_{d,n} \, |z_{d,n}, \beta_k)$.

In the LDA-type framework, both topic and word proportions (θ_d and β_k) are randomly chosen from a Dirichlet distribution with priors (the hyperparameters α and η). While such a standard topic model has proven to be quite efficient in discovering latent topics in economics and finance,[15] it has some limitations. First, the Dirichlet distribution does not allow topics within documents to be correlated and vary over time.[16] Second, the model does not permit topic prevalence and topic content to be influenced by exogenous factors or covariates. In other words, it does not allow us to discover topics and estimate their relationships with factors that may affect their dynamics. To overcome these limitations, we estimate OPEC topics and analyze their relationships with covariates using the Structural Topic Model (STM) developed by Roberts et al. (2013).

Both LDA and STM share the same spirit by estimating the quantities described previously. However, in STM, the estimation of the parameters depends on exogenous factors, X and Y (X and Y can be the same set of covariates). Technically, topic prevalence θ_d is assumed to be a random variable drawn from a Logistic-Normal distribution conditional on covariates, as:

$$\theta_d \mid X_{d\gamma}, \Sigma \sim Logistic - Normal(\mu = X_{d\gamma}, \Sigma)$$

where X_d is a vector of covariates, $\gamma \sim N(0, \sigma_k^2)$ is a matrix of coefficients with $\sigma_k^2 \sim Gamma(s^\gamma, r^\gamma)$, and Σ is the covariance matrix.

The topic content β_k is replaced with a multinomial logit such that a word's distribution is the combination of three effects (topic κ^k, covariates κ^y, and topic-covariate interaction $\kappa^{y,k}$) over v individual words in the relevant vocabulary of possible words, such as:

$$\beta_{d,k} \propto exp(m + \kappa_v^k + \kappa_v^y + \kappa_v^{y,k})$$

[15]See Hansen & McMahon (2016), Hansen et al. (2018), Larsen & Thorsrud (2019).
[16]See Blei & Lafferty (2006) and Blei & Lafferty (2007) for some extensions.

where m is the baseline word frequency, and $(\kappa_v^k + \kappa_v^y + \kappa_v^{y,k})$ is a collection of coefficients with $\kappa_v^{y,k} \sim Laplace(0, r_v^{y,k})$ and $r_v^{y,k} \sim Gamma(s^\kappa, r^\kappa)$.

This framework allows us to evaluate how our supply, demand, and speculative variables affect OPEC communication defined as topics and word proportions.

Measuring OPEC communication with mixed-membership topic models is difficult because of the latent structure of the parameters, as well as the intractable and non-convex posterior. Two approximate inference algorithms are popular for the estimation: Gibbs sampling (Griffiths & Steyvers (2004)) and variational inference (Blei et al. (2003)). As suggested by Roberts et al. (2016), we estimate the model using a semi-collapsed variational EM algorithm. We further induce sparsity on the collection of parameters by regularizing prior distributions for κ (with Laplace prior), and γ (with L1-penalty) to improve interpretability, prevent overfitting, and increase computational efficiency.[17]

A more challenging exercise in estimating topic models is the dimensionality of the latent space, namely the number of topics K. The procedure always involves a trade-off between statistical goodness-of-fit (i.e., higher K) and output interpretability (i.e., lower K).[18] We use different values of K, ranging from $K = 20$ to 60, and compute several statistical criteria (see Appendix A for technical details). We select $K = 40$ based on both statistical power and interpretability.

2.3 OPEC narratives and endogenous factors

2.3.1 Selecting topics from OPEC communication

As is common in text-mining, our OPEC press releases database is high-dimensional and sparse (a 262×12586 document-term matrix with 90% of scarcity). Consequently, we need to reduce the dimensionality of the corpus before estimation. In other words, we have to remove words containing little topical content (see Appendix A for technical details). In a nutshell, the process resides in removing stopwords[19] (i.e., *'the'*, *'are'*, *'but'*, ...), given names, surnames, numbers and punctuation, as well as converting remaining terms into their linguistic roots (i.e., stemming). Once the dimensionality problem is reduced, STM can be estimated on the new document-term matrix.

We estimate the 40-topics STM on our OPEC press release corpus from March 2002 to March 2021. The two main outputs are topics and word proportions covering different facets of OPEC communication. The model does not give any label, but provides the probability of each word within topics. While the label in itself plays no role in the analysis, it provides a convenient way to discuss OPEC communication. We propose to label topics based on both the top 10

[17] Additional results are available upon request from the authors.
[18] See Chang et al. (2009) for a discussion.
[19] The stopword list we used is from http://snowball.tartarus.org/algorithms/english/stop.txt, and is available upon request from the authors.

FREX (FRequency and EXclusivity) terms and most-probable bigrams (topic labeling and technical details are available in Appendix B).

We select some topics from the estimation to highlight different aspects of OPEC communication. They are represented as clouds of keywords in Figure 2 together with their labels. As shown, OPEC communications are very diverse, with topics related to crude oil prices (Topic 2, panel a) and shortages (Topic 3, panel b), production adjustments during turbulent times (Topic 20, panel e), economic growth (Topic 11, panel c), climate change (Topic 19, panel d), and energy policy (Topic 36, panel f).[20]

OPEC narratives over the whole period also reveal that some topics are related with others. In other words, specific topics in OPEC communication tend to co-occur during particular circumstances. Figure 3 provides a static picture of these correlations over the period as a communities network map using an infomap algorithm (see Rosvall & Bergstrom (2007) for more details).[21]

Several observations emerge from the map. First, OPEC has a large and well developed spectrum of communication with connected narratives. Second, the weight of each spectrum is not homogeneous with different levels of topics' importance: Topic 34 (OPEC/Non-OPEC Production Participation), Topic 25 (OPEC Production Adjustment), Topic 13 (Spare Oil Production Capacity), and Topic 4 (Oil uncertainty/Volatility) are examples of extensive topics. Third, while OPEC communication is very diverse, the overall narrative structure can be grouped into eight main communities in which topics are densely connected (each community is identified by a different color). These communities provide a better idea on the types of signals OPEC sends to markets. For instance, considering topics' labels, we can identify that the orange cluster is about "supply/production adjustment" signal during times of uncertainty, when cooperation within OPEC producers and between OPEC and non-OPEC countries is needed. While both dark blue and yellow groups seem to be related to long-term production, the light blue cluster appears to be linked to supply shortages linked to natural disasters. Similarly, the green cluster is primarily related to price volatility and crude oil market stability, and both purple and brown communities are associated to OPEC international relations.

Topics' evolution over time for each community is reported in Figures 13 to 18 in Appendix C. These figures provide insights on the link between market conditions and OPEC signals, and also help us understand the nature of the signals *per se*. For instance, Figure 13 clearly confirms the "supply adjustment" signal during periods of high uncertainty, such as the Afghanistan War in 2002, the US shale oil development in 2015-2018, and the COVID-19 pandemic. Figure 15 depicts "price volatility and market stability" signals which peak during periods of strong price fluctuations, for instance, during the GFC.

[20]For the full list of topics see Tables 6, 7, and 8 in Appendix B.

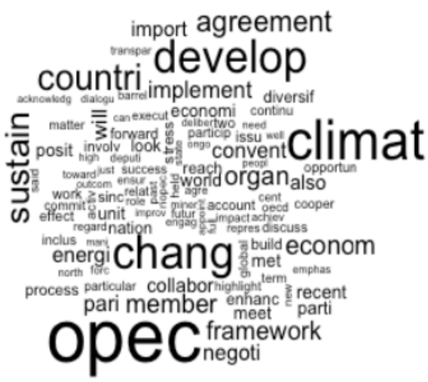
[21]As a robustness check (not reported), we also compute communities using walktrap (Pons & Latapy (2006)), louvain (Blondel et al. (2008)), and propagating labels (Raghavan et al. (2007)) algorithms, for which we get very similar clusters in terms of information variation (Meilă (2003)).

Figure 2: Selected topics from OPEC communication

(a) Topic label 2: Basket Price

(b) Topic label 3: Oil Shortage

(c) Topic label 11: Economic Growth



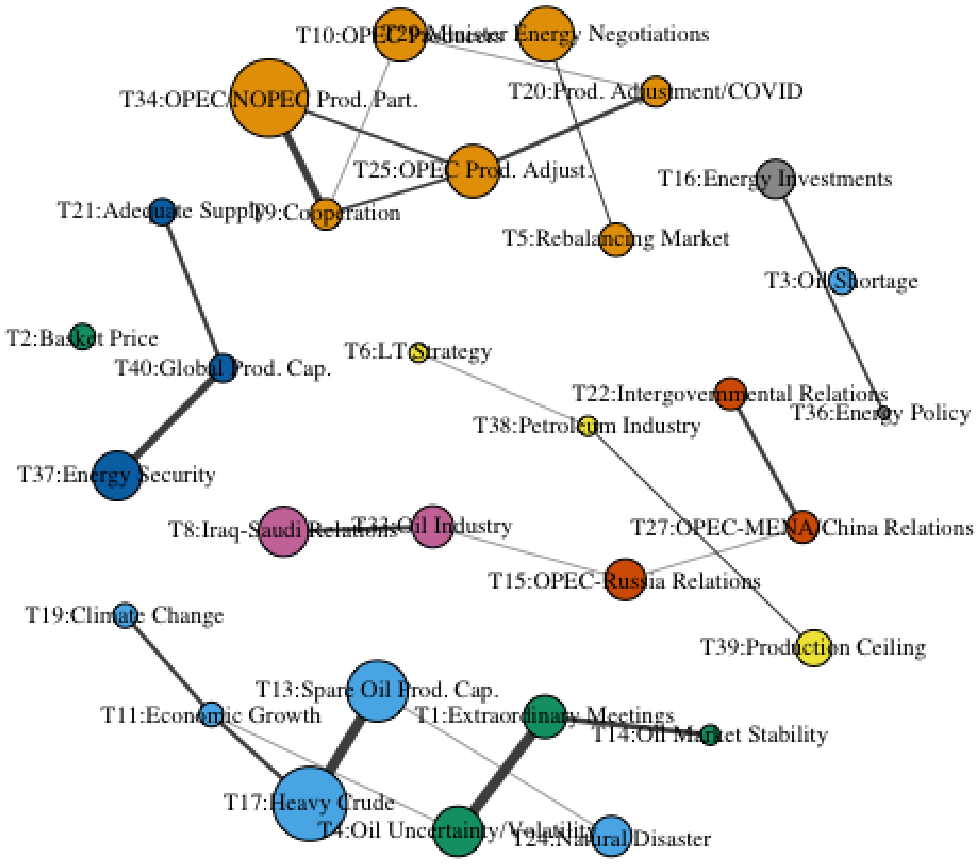
(d) Topic label 19: Climate Change

(e) Topic label 20: Prod. Adjustment/COVID

(f) Topic label 36: Energy Policy

Note: These figures report estimated topic distributions from STM as word clouds. The size of words in clouds corresponds to the probability of occurrence in the corresponding topic. The larger the word, the higher the probability to occur. Note that we report the stemmed tokens. The label is from the methodology discussed in Appendix B.

Figure 3: Communities topics network in OPEC communication



Note: This figure represents topic correlation over the whole period as a network structure. Node size indicates topic's weight in the whole system (i.e., the bigger the node, the more important is the topic with respect to others). Edge size (thickness) indicates the strength of the connection between two topics (i.e., the thicker the edge, the stronger the connection). Colors characterize nodes' communities based on an infomap algorithm. For simplicity, isolated nodes without huge contributions to the system have been removed from the network.

2.3.2 Drivers of OPEC communication

We are interested in identifying the variables that most impact OPEC topics. We extract from our STM model the estimated coefficients $(\gamma, \kappa_v^k, \kappa_v^y, \kappa_v^{y,k})$ and run linear regressions considering (i) each topic as endogenous, and (ii) supply, demand, and speculative side factors as exogenous. Measurement uncertainty, potential serial correlation, and heteroskedasticity problems are treated "locally" by stepping through each document, updating the parameters, then saving the local covariance matrix.

As shown in Figure 4, eight exogenous factors influence OPEC communication. Regarding the supply-side factors (in green), the impact of spare crude oil capacity on OPEC communication is particularly significant, especially when capacity is high. This result is expected since OPEC defines its production levels as a function of its reserves and demand conditions. Communicating on factors related to supply (Topics 21 and 37 in blue) when reserves are high is a way for OPEC to reassure markets about oil reserves in the event of strong demand, as well as about possible shortages (Topic 3 in light blue). In the event of high reserves, OPEC also communicates to stabilize the market (Topic 14 in dark green), for instance through production cuts to limit a potential price decrease. When reserves are low, communication mainly concerns cooperation between OPEC members as well as OPEC and non-OPEC countries.

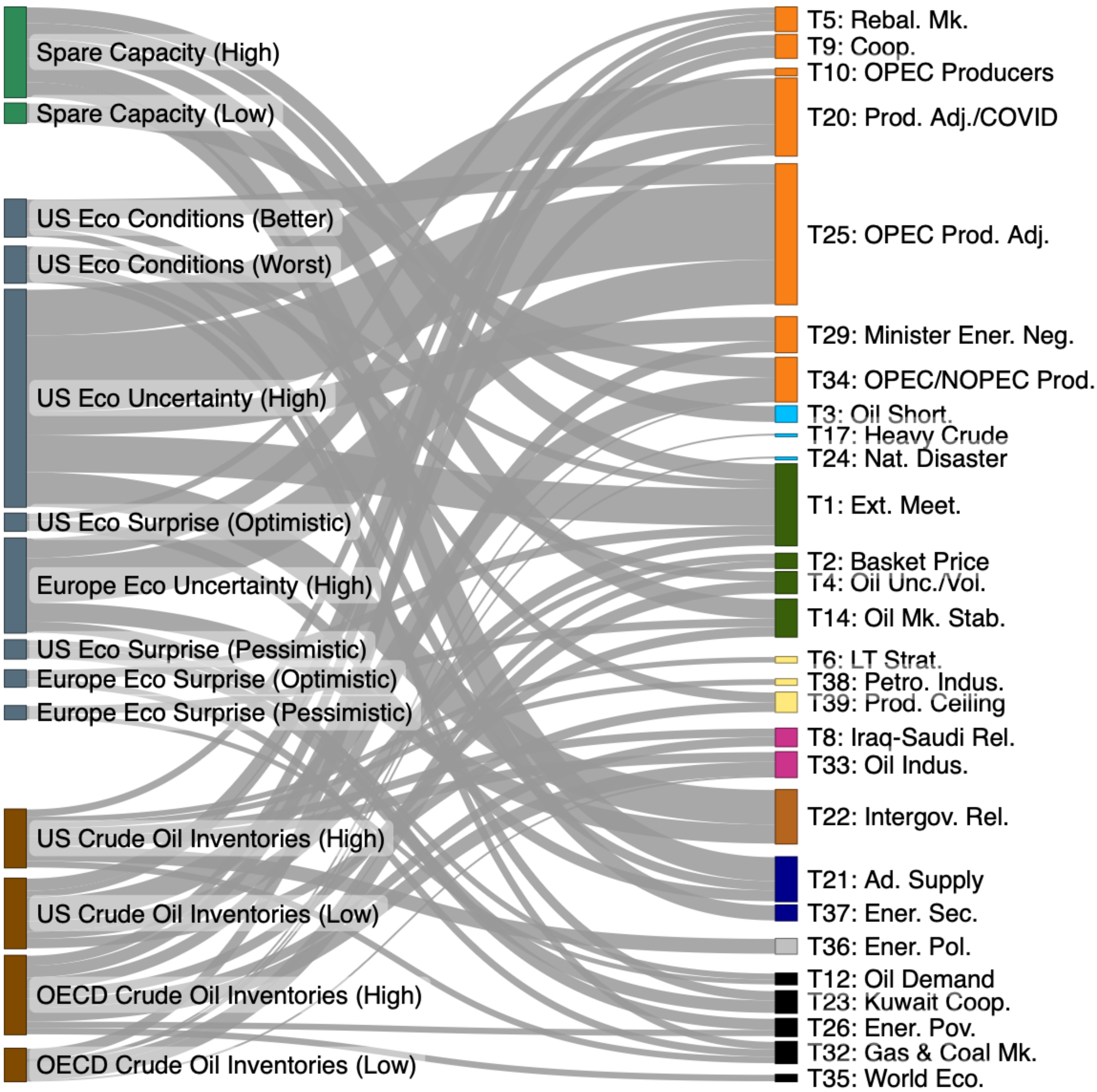
Turning to demand-side factors (in dark gray) high economic uncertainty in the US and, to a lesser extent, in Europe is highly significant in explaining OPEC communication. When economic conditions are highly uncertain, OPEC tends to intervene on topics related to production adjustments (Topics 20 and 25 in orange) and unscheduled extraordinary meetings (Topic 1 in dark green). These meetings are typically conveyed when unanticipated events occur and are usually associated with falling prices. OPEC communication aims at reducing uncertainty and reaffirming its market power to stop prices from falling further. As an illustration, the COVID-19 pandemic pushed OPEC to communicate its commitment to limit the negative consequences of the crisis on the oil market caused by the slowdown of economic activity due to the lockdown.

Similarly, a surprise effect regarding forecasters' predictions on the US and European economies, as well as the observed US economic conditions (in dark gray), affect OPEC communication. In the case of pessimistic surprises, OPEC mainly intervenes on demand-related concerns (Topic 12 in black), while it deals with topics related to its strategy (Topics 1 in dark green, and 5 in orange) when surprises are optimistic. In case of bad US economic conditions, OPEC communicates on supply-side factors to limit both volatility in the oil market and the subsequent fall in crude oil prices.

Speculative-related factors (in brown) affect several topics. OECD crude oil inventories significantly impact OPEC communication, mainly when stocks are high—i.e., outside the 5-year average band. In this case, OPEC communication focuses on several topics linked to its strategy (Topics 5 in orange, and 8 in purple), the oil market (Topics 14 in dark green, and 33 in purple), and the world economic situation (Topic 35 in black). OECD inventories have been relatively stable for many years, but an upward trend pushes OPEC communication to reaffirm its market power. A similar pattern is observed when US crude oil inventories are high. This is to be expected, since OPEC has a strong incentive to intervene when prices are relatively low.

In the case of low/normal inventories (in brown), OPEC communication mainly focuses on

Figure 4: Causes of OPEC communication



Note: This figure reports the effects of exogenous factors (left panel) on topics' distributions (right panel). Factors colors indicate supply (in green), demand (in gray), and speculative (in brown) covariates. Topics colors indicate communities as in Figure 3. Only statistically significant relationships (at the 5% level) are reported.

cooperation and agreements (Topics 1 in dark green, 9 and 34 in orange). The aim is to provide a reassuring speech, limiting market uncertainty and volatility (Topics 4 in dark green, and 39 in yellow) that could be caused by fears of insufficient stocks or even shortages. The underlying idea is that the credibility of OPEC signals' is higher when its members behave cooperatively. Indeed, when its members act in a non-cooperative way, geopolitical tensions are reignited, creating uncertainty and undermining OPEC's credibility.

3 Signalling game: causes and consequences

This section describes our methodology for testing the effectiveness of OPEC communication on both price volatility and trading positions. First, we briefly present the theoretical framework borrowed from Morris & Shin (2002) and Amato et al. (2002), taking the perspective of OPEC market power. Second, we delve into our empirical strategy and elaborate on the variables we have selected for analysis.

3.1 Theoretical framework

To illustrate the influence of OPEC communication on market volatility and trading behaviors, we introduce a simplified signal coordination game model, drawing inspiration from the central bank communication literature.[22] Like central banks, OPEC faces a dual challenge in communication: balancing market expectations with behavior coordination—reminiscent of Keynes' "beauty contest" analogy. OPEC grapples with trade-offs between transparency and signal credibility, especially in influencing expectations and pricing power in crude oil markets. The evolution of OPEC's role has been influenced by a shift towards futures markets, which encompass a heterogeneous pool of actors like producers, swap dealers, refiners, and money managers (Fattouh (2007)). Consequently, OPEC employs a multifaceted strategy combining production decisions and public communications to steer market expectations (Fattouh & Mahadeva (2013)). Despite these efforts, the efficacy of such tools remains under debate.[23] Numerous studies contend that OPEC's signals often lack credibility because they are cost-free and thus easily dismissed (Farrell & Rabin (1996), Fattouh (2007)). This skepticism leads markets to withhold judgment until OPEC's announced decisions are implemented. Contrary to this view, our paper argues that, when considering the full spectrum of OPEC communications, their public signals matter.

[22] For seminal works on this topic, see Morris & Shin (2002, 2005), Morris et al. (2006) as well as related discussions in Svensson (2006) and Ehrmann & Fratzscher (2007).

[23] For comprehensive discussions on production decisions and their effectiveness, see Fattouh (2007). For insights into the impact of OPEC communication on oil prices, consult Wirl & Kujundzic (2004), Demirer & Kutan (2010), and Brunetti et al. (2013).

Parallel to central banks, OPEC serves a dual role as (i) an observer, gathering clues to inform future actions, and (ii) a market influencer, shaping expectations.
In line with Morris & Shin (2002) and Ehrmann & Fratzscher (2007), we consider a continuum of private agents $i \epsilon [0, 1]$ who make decision p_i guided by a utility function:[24]

$$U_i (p_i, \theta) = - \left[(1 - r) (p_i - \theta)^2 + r \left(L_i - \bar{L} \right) \right] \tag{1}$$

where θ denotes market fundamentals (i.e., in our case, the determinants of the crude oil market), $r \in [0, 1]$ is a constant, $L_i = \int_0^1 (p_j - p_i)^2 \, dj$, and $\bar{L} = \int_0^1 L_j dj$. The first term on the right-hand side of this equation emphasizes that agents make decisions aligned with fundamentals, while the second term captures a coordination game between agents' actions, akin to Keynes's beauty contest.

The optimal decision rule of agent i is:

$$p_i = (1 - r) E_i (\theta) + r E_i (\bar{p}) \tag{2}$$

where $\bar{p}$ is the average decision across agents. The social planner aims to maximize welfare, which in our model is given by $W(p, \theta) = - \int_0^1 (p_i - \theta)^2 \, di$, focusing solely on the component related to market fundamentals.

Agents receive two types of signals about θ: an individual private signal specific to agent i, $x_i = \theta + \varepsilon_i$ and a public signal $y = \theta + \eta$. Upon receiving these signals, each agent's optimal decision rule simplifies to[25]

$$p_i = \frac{\alpha y + (1 - r) \beta x_i}{\alpha + (1 - r) \beta}. \tag{3}$$

leading to an expected social welfare function

$$V(\alpha) = - \frac{\alpha + \beta (1 - r)^2}{[\alpha + \beta (1 - r]^2}. \tag{4}$$

The crux of our investigation lies in assessing whether a more precise public signal (or increased transparency) positively or negatively affects welfare. According to Morris & Shin (2002), social welfare decreases with transparency if

$$\frac{\alpha}{\beta} < (2r - 1)(1 - r). \tag{5}$$

This counter-intuitive result serves as a cautionary note for authorities about the amount and type of information they disclose. Specifically in the OPEC context, this leads to an analysis of how OPEC could strategically decide on the depth of forward-looking communication when setting new production levels. The pivotal issue is to determine the threshold level of precision, denoted by $\bar{\alpha}$, that makes the release beneficial to societal welfare

$$\bar{\alpha} = \beta (2r - 1). \tag{6}$$

[24] We borrow the notations from Ehrmann & Fratzscher (2007).
[25] The expected value of the fundamentals is $E_i(\theta) = \frac{\alpha y + \beta x_i}{\alpha + \beta}$.

This suggests that public signals with a precision level of α are more likely to be advantageous when the private signals are of lower quality—specifically, when $\alpha > \beta$. The crux is whether the precision of the public information is sufficient to warrant its disclosure.

Based on this theoretical groundwork and Equation (6), we set the stage for two important questions:

1. Does OPEC signal matter on average?
2. How does OPEC signal change with respect to the quality of the private signal?

These questions lead to empirically testable hypotheses, which we explore in the next subsection. In particular, we introduce an empirical framework specifically designed to scrutinize the effectiveness of OPEC' signalling – as gauged by OPEC communications – in shaping market expectations. Our empirical framework also addresses the dilemma OPEC faces over the level of information it should disclose about its market perceptions.

3.2 Empirical strategy

As discussed in the previous subsection and in Ehrmann & Fratzscher (2007), the Morris-Shin model postulates that both the public and private signals have an impact on social welfare. To measure the effectiveness of the OPEC signal on the crude oil market, we consider a two-step procedure that consists of (i) estimating the OPEC signal using the unsupervised learning model described in Section 2.2, and (ii) measuring and testing the effect of signals on the crude oil market using penalized regression. Let us present the econometric framework.

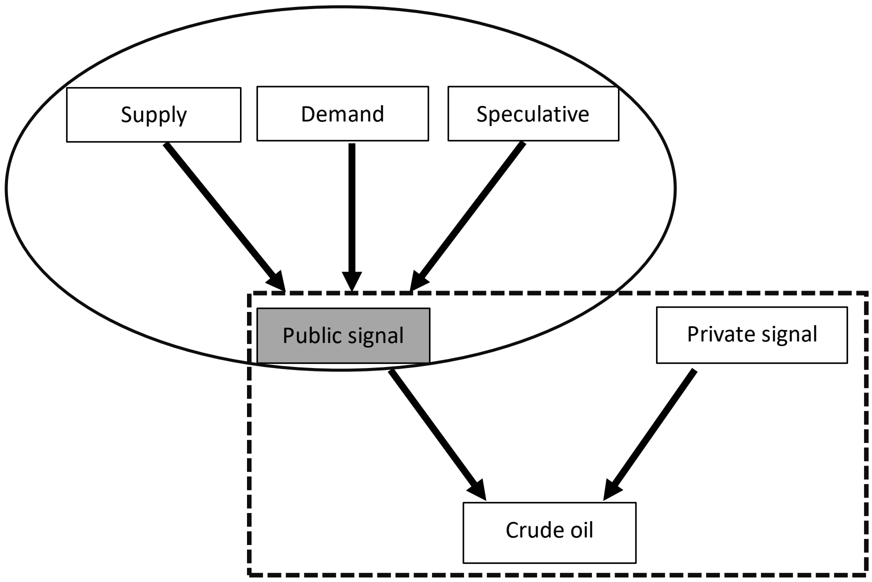
To test the effectiveness of OPEC signal on the crude oil market, we rely on two empirical models. The first specification is given by:

$$y_t = \kappa + \lambda z_{\alpha,t} + \mu z_{\beta,t} + \vartheta X_t + \upsilon_t \tag{7}$$

where y_t stands for the endogenous variable, κ is the constant term, X_t denotes a set of control variables, and λ and μ capture the effect of public and private signals on the crude oil market, respectively. Statistically significant estimates of λ and μ imply that both private and public signals convey important information. The public signal, however, should have a larger impact since it is related to underlying fundamentals as in Morris et al. (2006) and speaks to the effectiveness of OPEC communication.

We are aware that OPEC communications are endogenous to market fundamentals.[26] In other words, based on market fundamentals, OPEC decides to communicate the level of intensity of the public signal to send. OPEC can then act either as a catalyst or a buffer of market fundamentals to influence prices and expectations. The intuition of Equation (7) is illustrated in Figure 5. The circle frame indicates the machine learning part of our approach devoted to estimating OPEC's public signal (see Section 2.2). The dashed frame is the econometric part of the model, which we use to test the implications of the Morris et al. (2006)'s model.

Figure 5: Empirical model 1



Note: This figure reports our empirical strategy for Equation (7). The shaded variable is estimated out of the econometric model. Unshaded variables are observed. The circle frame is the machine learning part devoted to the estimation of the OPEC signal as described in Section 2.2. The dashed frame is the econometric part of the model. For simplicity, this figure omits control variables.

Equation (7) and Figure 5 imply that market participants directly observe both private and public signals (dashed frame). Yet, OPEC's public signal reflects the Organization's perception of crude oil market conditions, such as supply, demand, and speculative components (circle frame).[27]

The second empirical model investigates the public signal effect on the crude oil market depending on the precision of the private signal. In other words, we evaluate whether private

[26]As we will describe further, to deal with such endogeneity issues and to precisely identify the effect of the public signal, our experiment design is strictly limited to the days (weeks) when OPEC communicates. By considering the specific window of the announcements, we are able to isolate the impact of OPEC narratives on the market variables we are analyzing.

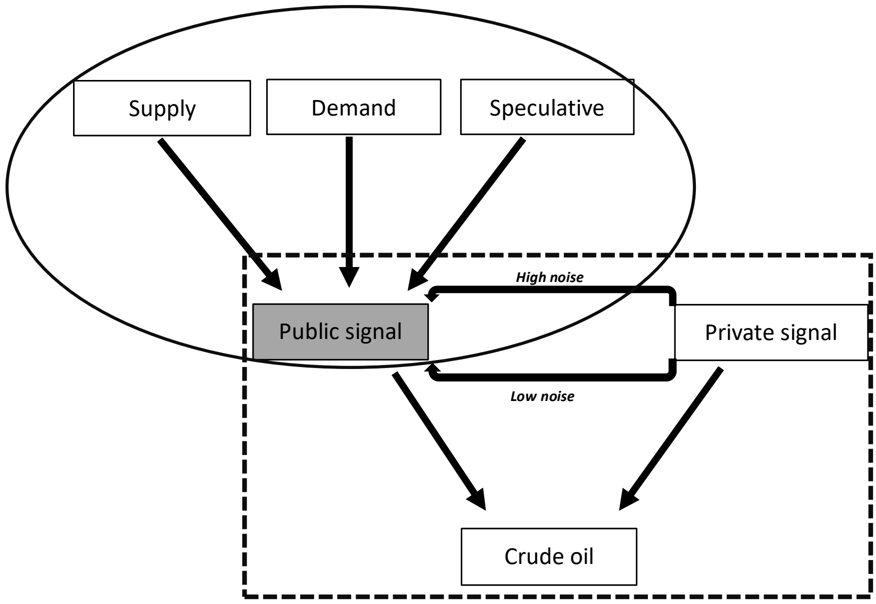
[27]See Kilian & Murphy (2014), Baumeister & Kilian (2016), and Brunetti et al. (2016) for a discussion on oil market determinants.

signal noise affects the information flow of the public one (as discussed by Morris et al. (2006)). To this end, we build a dummy variable $D^{\beta} = 1$ if the noisiness of private information is above its mean over the whole sample period (high noise) and 0 otherwise (low noise). We then estimate the following equation which includes an interaction term between the public signal and noisiness in private signal

$$y_t = \kappa + \lambda_1 \left(z_{\alpha,t} D_t^{\beta} \right) + \lambda_2 \left(z_{\alpha,t} \left(1 - D_t^{\beta} \right) \right) + \mu z_{\beta,t} + \vartheta X_t + \upsilon_t \tag{8}$$

where λ_1 (λ_2) denotes the effect of public signal in periods of high noise (low noise) in private information. κ, μ, and ϑ still stand for the constant, the effect of the private signal, and the control variables. The intuition of Equation (8) is illustrated by Figure 6, which is similar to Figure 5 with the addition of the noisiness of the private signal.

Figure 6: Empirical model 2



Note: This figure reports our empirical strategy for Equation (8). The shaded variable is estimated out of the model. Unshaded variables are observed. The circle frame is the machine learning part devoted to the estimation of the OPEC signal as described in Section 2.2. The dashed frame is the econometric part of the model. For simplicity, this figure omits control variables.

Equation (8) relates the effectiveness of public information to the quality of the private one. More generally, Equations (7) and (8) act as the empirical counterparts of the theoretical expressions (4) and (6) in Morris-Shin's framework. In particular, λ and μ in the empirical setting (Equation (7)) stand for α and β in the theoretical one (Equation (4)). Additionally,

λ_1 in Equation (8) captures the $\alpha > \beta$ condition implied by Equation (6).

3.3 Variables selection

3.3.1 Dependent variables

Social welfare is difficult to capture in our framework. Hence, we assess the effectiveness of OPEC signal on oil market dynamics using two categories of dependent variables: (i) crude oil futures price volatility, and (ii) trading positions of crude oil futures market participants. Our analysis spans two time periods: March 2002-March 2021 and June 2006-March 2021, based on data availability.[28] Traditional finance models argue that both volatility and trading volume serve as channels for new information to enter markets.[29] Moreover, the stated objective of OPEC is to coordinate the policies of member countries to stabilize the oil market. Therefore, it is natural to consider oil price volatility when assessing the efficacy of OPEC's public signal. Similarly, looking at market participants' positions guides our understanding of the functioning of the crude oil market and provides evidence of the reaction of market forces to OPEC communication.

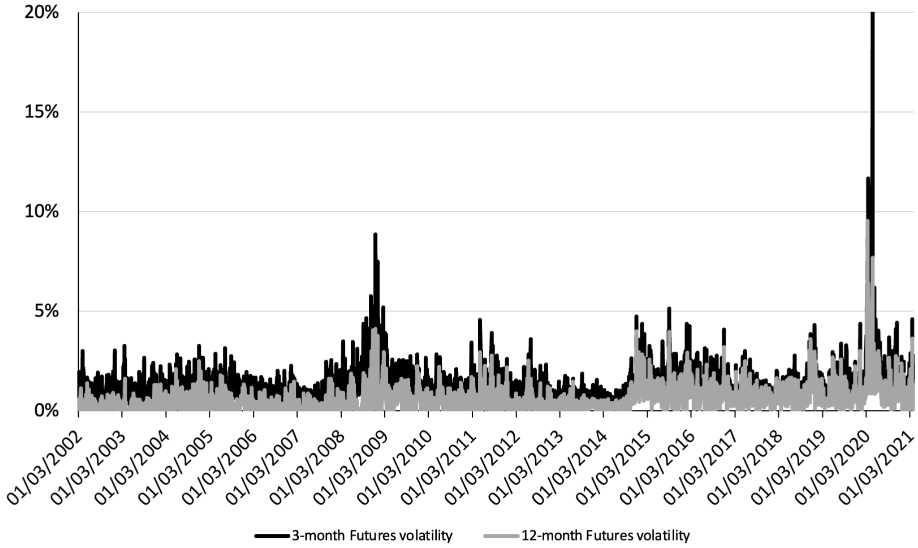
We employ daily volatility estimates from WTI futures contracts traded on the Chicago Mercantile Exchange over the 2002-2021 period. Volatility is measured by the daily range—the difference between log-high and log-low prices, a widely-accepted estimator of price volatility.[30] Our study incorporates the entire term structure of futures prices, ranging from 1- to 12-month maturities. This breadth allows us to understand how OPEC communications may differentially affect price expectations at various time horizons. Figure 7 reports the 3- and 12-month futures price volatility. As expected, volatility is time-varying, and shorter-maturity contracts exhibit higher volatility than longer-maturity contracts. Two notable periods can be distinguished with important spikes: the Global Financial Crisis and the COVID-19 pandemic.

[28]To investigate the sensitivity of our results to crisis episodes, such as the COVID-19 pandemic, we also assess OPEC communication effectiveness during those specific periods (see Section 5.1.1).

[29]See, Epps & Epps (1976), Tauchen & Pitts (1983), and Gallant et al. (1992).

[30]See, for instance, Brunetti & Lildholdt (2007). For robustness, we also consider squared returns as an alternative measure of volatility.

Figure 7: WTI crude oil futures volatility (2002-2021)



Note: This figure reports the daily WTI crude oil futures price volatility for the 3-month and 12-month maturity contracts. Volatility is proxied by log(high) - log(low) price. Data on WTI futures contracts are taken from the Chicago Mercantile Exchange.

We consider weekly trading positions as measured by the DCOT from the CFTC over the 2006-2021 period.[31] All the DCOT reports provide a breakdown of each Tuesday's open interest for markets in which at least 20 traders hold positions equal to or above the reporting levels established by the CFTC. We concentrate on four categories of traders as classified by the CFTC: traditional hedgers, swap dealers, money managers, and other reportable traders.

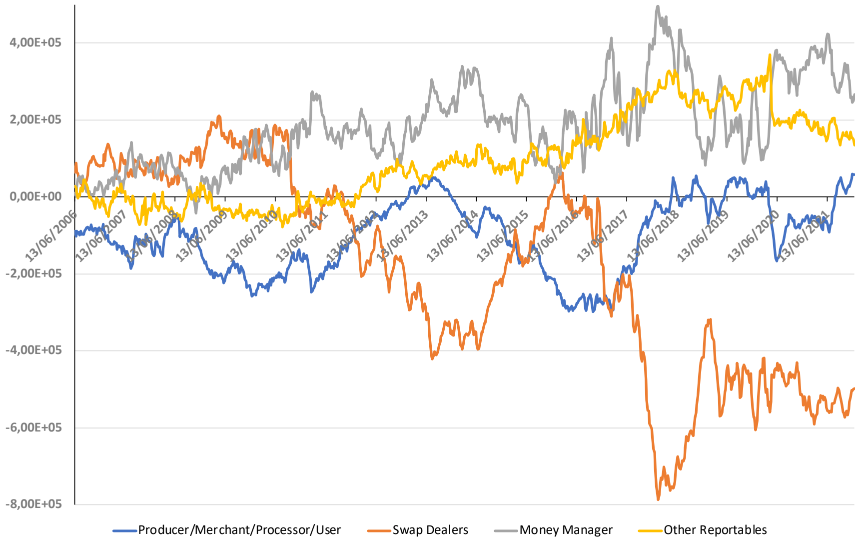
Traditional hedgers are producers, merchants and dealers (i.e., wholesalers, exporters, importers, shippers, etc.), as well as processors and users (i.e., fabricators, refiners, etc.)—hereafter P/M/D/P/U. Their main line of business concerns the physical commodity, as they are primarily engaged in the production, processing, packing, or handling a physical commodity, and they use futures markets to manage or hedge risks associated with their main activities. We also consider the positions of financial participants, such as swap dealers (SD) and money managers (MM). The former use futures markets to manage or hedge the risk associated with swap transactions. The swap dealers' counterparties may be speculative dealers, like hedge funds, or traditional commercial clients that manage risk arising from their dealings in the physical commodity. Money managers are engaged in managing and conducting futures trading on behalf of clients (i.e., registered commodity trading advisory, registered commodity pool operators, or unregistered funds identified by the CFTC). Finally, the last residual category consists of positions of all other reportable (OR) traders not included in the previous

[31]DCOT reports are weekly publications showing holdings of different participants in futures markets. The reports are published by the CFTC and contain details on long and short open interest positions of selected categories of market participants in each futures market. See, among others, Sanders & Irwin (2013).

three categories.[32]

For each trader category, we compute the weekly net positions as the difference between long and short positions as reported in Figure 8. P/M/D/P/U are on average short, which is in line with their role of hedgers—e.g., a producer sells its production in advance in the futures market (short position) to reduce price uncertainty. Money managers are usually long and may act as counterparties to hedgers. Swap dealers' net positions change dramatically over time depending on their swap business, but they are usually short.

Figure 8: Net disaggregated trading positions (2006-2021)



Note: This figure depicts the weekly net trading positions (source: DCOT reports from the CFTC) as measured by the difference between long and short positions for each considered category of traders.

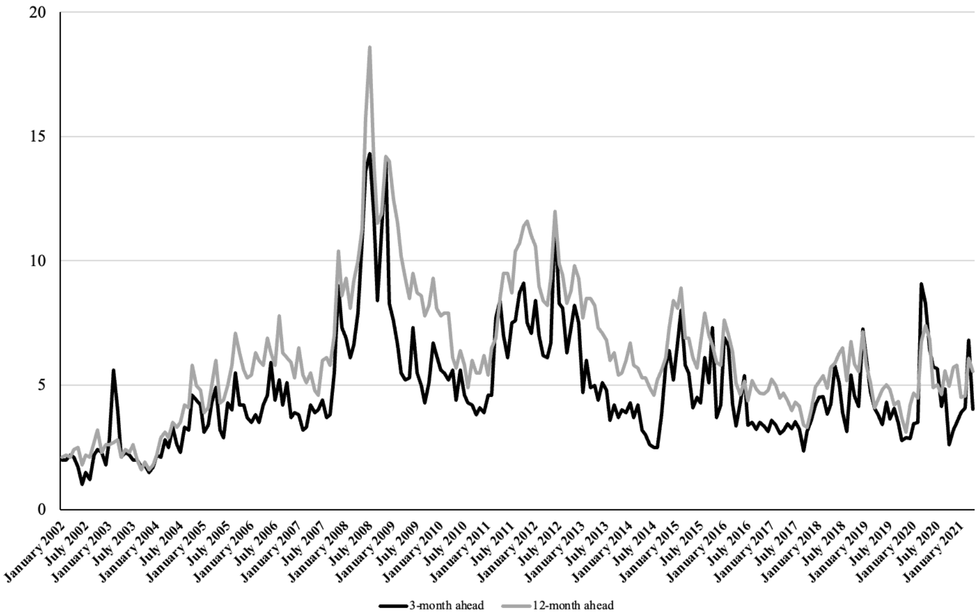
3.3.2 Independent variables

We employ two types of signals: public and private. The public signal consists of OPEC communications, quantified through textual analysis, as detailed in Section 2. The private signal is approximated by the Consensus Forecast Inc., which conducts monthly surveys on oil price forecasts of about 30 market participants for 3-month and 12-month horizons. While these forecasts are certainly influenced by crude oil market fundamentals, differences across agents should reflect private information. Hence, the cross-sectional standard deviation across experts serves as our measure of the private signal—see Figure 9.[33]

The private signal is very high during the Global Financial Crisis (GFC), in 2011-12 when oil prices went above $100 per barrel and there was political instability in some producing

[32] See https://www.cftc.gov/MarketReports/CommitmentsofTraders/index.htm for more details.
[33] See, Ehrmann & Fratzscher (2007) for further discussion.

Figure 9: Standard deviation of WTI oil price consensus forecasts (2002-2021)



Note: This figure reports the standard deviation of the consensus forecasts on the oil prices for 3-month and 12-month horizons.

countries, in 2015-16 when oil prices plunged driven by a growing supply glut, and at the beginning of the COVID-19 pandemic. Overall, the private signal captures significant market developments.[34]

To capture significant market events, we control for various external factors connected to OPEC's activities.[35] Dummy variables for the March 2002-March 2021 period are introduced as follows:

- Production decision variables: (i) $D_t^{p_1} = 1$ else 0 when OPEC decides to increase production, and (ii) $D_t^{p_2} = 1$ else 0, when OPEC decides to decrease production. The baseline case is "neutral" when either the level of production is kept unchanged or when there is no mention of any other decision.

- Meeting type variable (anticipated vs. unanticipated announcements): $D_t^m = 1$ else 0, when the OPEC meeting is not scheduled in advance (e.g., ordinary vs. extraordinary meetings).

- OPEC behavior (acting as a cartel or not): $D_t^b = 1$ else 0, when OPEC members cooperate during the period.[36]

[34]See Ehrmann & Fratzscher (2007).

[35]For robustness, we also consider the global macroeconomic environment; see Section 5.

[36]A cartel is defined as production coordination with respect to quotas (see Brémond et al. (2012) for an empirical analysis). To distinguish periods of cooperative and non-cooperative OPEC behaviors, we follow the

4 Cheap talk or credible signal?

Our identification strategy is similar, in spirit, to Känzig (2021) and is based on two considerations. The first relates to the dominant role of OPEC in the crude oil market. OPEC produces about 40 percent of the world's crude oil and accounts for an estimated four-fifths of total crude oil reserves. Hence, it is a major player in the crude oil market. Market participants pay attention to OPEC announcements, which can be considered the dominant event on the days when they are pronounced. The second is related to our experiment design, which is limited precisely to the days when OPEC announcements occur. Considering the specific window of the announcements' days allows us to isolate the impact of OPEC narratives on the market variables we analyze.

Applying this methodology to estimate the effects of OPEC communication on the volatility process is straightforward since we have access to daily volatility data. It is more complex when considering traders' positions which are measured at a weekly frequency (they are reported every Tuesday). To overcome the problem of mismatch frequencies between OPEC announcements and trading positions, we align data points with respect to the OPEC signal by considering either the corresponding day of the announcement (when it coincides) or the closest, but preceding, available day. Figure 10 depicts two examples. First (in green), both the signal and the positions are reported on the same day (Tuesday), and no alignment is needed. Second (in red), OPEC signal occurs on Friday (Week 1 - Friday) between two reported trading positions (Week 1 - Tuesday and Week 2 - Tuesday). Because traders do not know in advance the content of OPEC communication, we align the signal (Week 1 - Friday) to the next available data point (Week 2 - Tuesday). It is important to note that our alignment strategy may introduce a downward bias to our estimates since trader positions may quickly react (within the same day) to OPEC communications.

We also acknowledge that other factors may drive both OPEC communication and positions of market participants in a given week.[37] To address this issue, we consider various fundamental determinants of the public signal, allowing us to account for such factors that may affect OPEC announcements.[38] Overall, we believe our approach is able to identify the effects of OPEC narratives on volatility and traders' positions.

In order to estimate Equations (7) and (8) and test the credibility of OPEC signal, we rely on Lasso penalized regressions. The reason is twofold. First, technically our framework faces a dimensionality problem as the sample size is not large enough compared to the parameters'

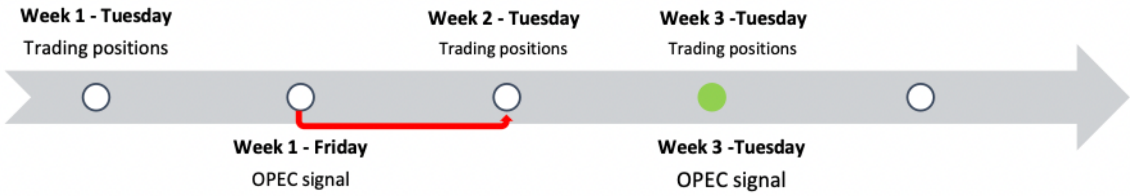
methodology discussed in Almoguera et al. (2011) and compare production quotas assigned by OPEC to the actual production levels. If actual production in period t is at least 5 percent over the quota established for that period, it indicates non-cooperation otherwise cooperation.

[37] Ideally, we would like to have access to daily traders' positions.

[38] Those variables are described in Section 2.1. Note that our results remain robust to various specifications (see Section 5 and Appendix E).

space. Equations (7) and (8) count 46 and 86 coefficients, respectively, while our sample consists only of 262 observations. Second, analytically the purpose is to investigate which set of signals is credible for price volatility and trading positions. Lasso regressions allow us to overcome both problems by selecting variables that are statistically relevant, and forcing to zero the coefficient of less important variables.[39]

Figure 10: Frequency mismatch alignment



As common for penalized regressions, the constraints on the size of coefficients depend on the magnitude of each variable. Therefore, as recommended by Tibshirani (1997), we standardize all our variables.[40] Table 1 displays the estimated coefficients of the Lasso regression for Equation (7). We report results for price volatility at 1-, 6- and 12-month maturities, and trading positions for the considered categories.[41]

From a general viewpoint, Table 1 shows that OPEC communication affects traders' positions more than price volatility. For each category of factors, several topics have a significant impact on traders' positions. This is particularly true for swap dealers (SD), money managers (MM), and other reportable (OR) traders. It is worth noting the dichotomy between traders engaged in the physical business (Producers/Merchants/Processors/Users) and financial traders (SD, MM, OR). Topics which are statistically important for the former are not relevant for the latter categories of traders. The only exception is Topic 9, cooperation, which implies that market participants pay attention to the credibility of OPEC communications as captured by cooperation. The effect of OPEC communication on traders' positions is thus dependent on whether they are involved in physical or financial activities, and is particularly significant for swap dealers. This result can be explained by the fact that SD rebalance their portfolio frequently over time. They are particularly sensitive to OPEC signals aiming at stabilizing the market, as they are mainly interested in managing and hedging the risk associated with swap transactions.

[39]For recent applications of penalized Lasso regressions in finance, see Chinco et al. (2019), Calomiris & Mamaysky (2019), Kozak et al. (2020), Freyberger et al. (2020), and Gu et al. (2020).

[40]We also standardize dummy variables. Interpretation of standardized dummy variables in penalized regressions is often difficult. So, as a robustness check, we also perform regressions with non-standardized dummy variables (see Section 5). Results are similar to those reported in the paper and are available upon request from the authors.

[41]The results for other maturities are available upon request from the authors.

Regarding oil price volatility, OPEC tends to reduce it significantly when it intervenes on topics related to uncertainty and volatility (Topic 4), global production capacity (Topic 40), and energy policy (Topic 36). In those cases, the public signal is credible in that OPEC's reassuring communication contributes to stabilizing the oil market. It is worth mentioning that credibility increases with maturity as some topics—such as those related to long-term production and the petroleum industry—become significant at a 12-month horizon. This result is highly interesting since it shows that OPEC communication effectively reduces oil price volatility and favors market stability, especially for long-term contracts, even in the presence of significant private signals. These results contrast those in Demirer & Kutan (2010). However, the technical approaches and the relevant variables are different.[42]

Overall, Equation (7) allows us to explain a substantial part of the fluctuations of price volatility and trading positions over the considered period. Based on the adjusted-R^2, our results show that, on average, 34 percent of crude oil price volatility variation and 57 percent of net trading positions variance are explained by our model.[43] For P/M/U/P, SD, and OR we get more than 50 percent of explanatory power.

Tables 2 and 3 report the effect of OPEC signals on both price volatility and trading positions depending on the level of the noise in private signals (at 3 months)—Equation (8).[44] The intuition is that OPEC communication may gain credibility when private signals become uncertain.

As shown in Table 2, the effect of OPEC communication on price volatility appears to be more credible when private signal noise is high (credibility is measured as the proportion of significant topics in low and high regimes). OPEC credibility is, however, not constant and varies with maturity (four topics are significant at 1-month against twelve and eleven for 6- and 12-month, respectively). Based on the adjusted-R^2, OPEC communication is an important element of price volatility, especially in longer maturity contracts (on average, the regressions explain 48 percent of the variance of the endogenous variables). Combining the results in Tables 1 and 2, we find that the maturity of the crude oil futures contract matters. In particular, our results indicate that OPEC signals have a stronger impact on longer maturity contracts. In fact, while only supply-related topics (Topics 20 and 29) are important at 1-month horizon, multiple types of signals on supply, price, and shortage are significant at 6- and 12-month maturities—the most important are global supply (Topic 20), market stability (Topic 14), economic growth and oil demand (Topics 9 and 12). Multiple signal communication is then an efficient strategy in the long run if the Organization intends to affect price volatility.

[42]Demirer & Kutan (2010) rely on the event study methodology to assess the effects of OPEC conferences on the crude oil market activity in the US.

[43]We also use the deviance ratio as a measure of the explanatory power of our model. Results are robust and available upon request from the authors.

[44]Results for 12-month noise are similar to those reported and are available upon request from the authors.

Table 1: Effectiveness of public and private signals

Signals	1-month	6-month	12-month	P/M/D/P/U	MM	SD	OR
Public signals							
Supply adju. signal							
T5: Rebalancing Market	X	0.029	0.023	-7.403*	X	11.16	X
T9: Cooperation	X	0.054*	0.033	15.83*	X	-32.72*	16.20*
T10: OPEC Producers	X	X	0.0052	5.145	X	-9.743	7.371*
T20: Prod Adjustment/COVID	0.153*	0.043	0.035	3.647	X	-23.46*	16.00*
T25: OPEC Prod. Adjust.	X	X	X	2.377	18.35*	-26.95*	1.346
T29: Minister Energy Negotiations	0.090*	0.017	0.016	-6.875	-5.415	23.42*	-1.999
T34: OPEC/NOPEC Prod. Part.	-0.027	-0.013	X	4.243	21.54*	-43.75*	19.55*
OPEC internal. relations signal							
T15: OPEC-Russia Relations	X	X	X	12.62*	X	X	X
T33: Oil Industry	X	X	X	-7.955	14.77*	X	X
Price vol. & Mk stability signal							
T1: Extraodinary Meetings	X	-0.021	X	-2.485	-9.894*	27.87*	-6.334
T4: Oil Uncertainty/Volatility	-0.043*	-0.042*	-0.069*	X	-15.17*	45.76*	-20.77*
T14: Oil Mkt Stability	0.342*	0.179*	0.119*	-7.131*	-2.202	5.662	X
Prod. Ceiling & LT inv. signal							
T6: LT Strategy	X	X	-0.0025	-1.014	-2.162	16.23*	-8.490*
T21: Adequate Supply	X	X	-0.061*	X	-13.41*	29.08*	-11.30*
T38: Petroleum Industry	X	X	-0.031*	-3.634	-9.307*	29.85*	-8.688*
T40: Global Prod. Capacity	-0.040*	-0.065*	-0.033*	X	-6.961*	18.49*	-10.05*
Oil shortage signal							
T11: Economic Growth	X	0.027	0.062*	14.72*	X	X	-3.013
Energy policy signal							
T36: Energy Policy	-0.050*	-0.056*	-0.021*	-4.681	-9.609*	42.66*	-19.07*
Others signal							
T7: Physical/Financial Interaction	-0.016	-0.023	-0.012	8.083*	X	X	-1.329
T23: Kuwait Cooperation	X	X	X	-2.506	8.605*	X	1.885
T30: Energy Outlook	X	X	X	1.837	14.95*	-19.58*	X
T35: World Economy	X	X	X	-8.188*	-1.468	30.08*	-15.13*
Private signals							
Consensus 3 months (SD)	0.292*	0.126*	0.053*	29.07*	-5.832	X	15.66*
Consensus 12 months (SD)	X	X	X	X	-20.00*	38.43*	-53.31*
Control variables							
Production Decision: Cut	2.143*	0.150*	0.150*	-2.978	X	X	X
Production Decision: Increase	X	0.014	X	X	X	X	-3.850
Extraordinary Meetings	X	X	X	X	X	X	X
OPEC Cooperation	X	0.049*	0.047*	-48.16*	X	48.98*	-8.398*
Adjusted R^2	37%	33%	33%	56%	44%	58%	68%

Note: This table reports the estimated coefficients of the penalized bootstrap Lasso regressions for both price volatility (in percentage points) and trading positions organized as clusters (see communities in Figure 3). We only report statistically significant topics. $*$ indicates significance at the 5% level. "X" indicates zero value coefficients. To save space, coefficients of trading positions are divided by 1000.

The effect of OPEC communication on trading positions (Table 3) is generally more significant than on price volatility. Our models explain on average 61 percent of net positions' variability across all traders. The adjusted-R^2 amounts to more than 50 percent for P/M/D/P/U (67 percent), SD (59 percent), and OR (67 percent). While OPEC signal matters for each trading category, it is even more important when private noise is high and mainly for investors that are predominantly engaged in physical commodities (namely P/M/D/P/U). Surprisingly, for traders involved in financial activities with no physical exposures (MM and SD), OPEC credibility is qualitatively unaffected by the amount of private noise. Similar to the results in Table 1, the dichotomy between financial and physical traders also matters in terms of the signal (topics of OPEC communication). Physical traders' (P/M/D/P/U) net positions are mainly positively affected by supply- (Topics 9 and 34) and shortage-related (Topic 11) topics. Financial (MM and SD) and OR net positions, on the other hand, are impacted by many topics. In particular, MM and OR mainly respond positively to supply-related communications and negatively to long-term investment signals. SD move in the opposite direction, reacting negatively to supply signals and positively to price and investment signals, perhaps because SD act as a counterpart to MM and OR. Interestingly, in line with current debates, Topic 19 on climate change only impacts positions from traders engaged in physical commodities (P/M/D/P/U) but has no role on financial ones. It is also important to point out the strong effect of OPEC cooperation on trading positions.

Overall, our results indicate that OPEC communication is relevant and effective. First and importantly, it achieves the objective of stabilizing the crude oil market by reducing volatility levels of crude oil futures prices (Table 1). These results are stronger for longer maturity contracts, indicating that OPEC communication affects the term structure of oil futures. Since the main mandate of the Organization is to stabilize the oil market, our results provide evidence that OPEC fulfills its mandate. Second, market participants react to OPEC communication by readjusting their net positions. This represents additional evidence that OPEC communication matters. Topics covered in OPEC communications explain a large part (measured by adjusted-R^2) of the variation in trader positions. Topics related to OPEC credibility seem to be particularly relevant. Finally, our results represent an empirical test of the Morris and Shin (2002) theory. OPEC communication seems to provide a credible signal, and this signal is stronger the higher the noise in the private signal.

5 Robustness and placebo tests

5.1 Robustness checks

5.1.1 OPEC communication in recent crises

As emphasized above, OPEC communications are credible in providing signals about market fundamentals influencing crude oil price volatility and net trading positions. Credibility increases when public announcements interact with noisy private signals, making communications an important tool to shape expectations.

During financial and economic crises, however, OPEC lacks the ability to implement efficient coordinated production decisions. This is mainly due to the structure of the Organization, which lacks a formal enforcement mechanism constraining members to comply with the agreed production quotas.[45] As for central banks, public communications then become even more important to anchor market expectations.[46] To investigate how efficient OPEC communications are during unconventional times, we consider two recent turmoil periods included in our sample: the Global Financial Crisis (3/5/2007 - 12/17/2008) and the COVID-19 pandemic (3/5/2020 - 3/4/2021).[47] We gather these two episodes together to avoid small sample issues, and run our estimations of Equations (7) and (8).

Table 4 reports the effectiveness of OPEC communication during crisis times. Interestingly, compared to our previous results (see Table 1), the explanatory power (based on the adjusted-R^2) increases substantially for all of our dependent variables, but P/M/D/P/U (yet, it remains high at 50 percent). However, only a few topics are statistically significant. Not surprisingly, "OPEC cooperation," "Production adjustment/COVID," and "Oil market stability" are important topics for the volatility process. The coefficients are positive, which implies that OPEC communication relative to those topics is associated with higher volatility, perhaps linked to OPEC's fragile structure in terms of enforcing production decisions. Positions of both commercial and non-commercial traders are mainly characterized by negative coefficients, implying that OPEC communications are associated to a reduction in net traders' positions. The opposite is true for other positions. Interestingly, Topic 15 refers to the OPEC-Russia relationship and is important for traders' positions, suggesting that in crisis periods, OPEC+ alliance may play a critical role.

Tables 9 and 10 in Appendix D report the results from Equation (8) with high private noise over the selected crisis periods. Compared to Tables 2 and 3, the explanatory power of our models is globally more important during crisis periods for both crude oil price volatility and trading positions. In line with the Morris-Shin predictions and with our findings in Section 4, OPEC's public signal is more significant when private noise is high.

[45] See Fattouh & Mahadeva (2013).

[46] See Eggertsson & Woodford (2003), Coenen et al. (2017), and Hubert & Labondance (2018) for some discussions on the role of central banks' communications during unconventional times.

[47] Selected time ranges are based on NBER business cycles dating.

Table 2: Effectiveness of public and private signals on price volatility in high and low private noise (3 months)

Signals	1-month		6-month		12-month	
Public signals	*High*	*Low*	*High*	*Low*	*High*	*Low*
Supply adju. signal						
T9: Cooperation	X	X	0.024	0.066*	0.004	0.062*
T20: Prod Adjustment/COVID	0.228*	X	0.072*	X	0.069*	X
T25: OPEC Prod. Adjust.	X	X	0.030	X	0.049*	X
T29: Minister Energy Negotiations	0.061*	0.023	-0.028*	0.064*	X	0.046*
OPEC internal. relations signal						
T8: Iraq-Saudi Relations	0.130*	X	0.107*	X	0.126*	X
Price vol. & Mk stability signal						
T1: Extraodinary Meetings	X	X	-0.022*	X	X	X
T4: Oil Uncertainty/Volatility	X	X	X	-0.025	-0.029*	-0.025*
T14: Oil Mkt Stability	0.401	X	0.218*	X	0.149*	X
Prod.Ceiling & LT inv. signal						
T21: Adequate Supply	X	X	-0.016	X	-0.010	-0.054*
T38: Petroleum Industry	0.028	X	0.043*	-0.041	-0.008	-0.020
T40: Global Prod. Capacity	X	-0.0092	-0.027*	-0.069*	X	-0.039*
Oil shortage signal						
T11: Economic Growth	X	X	0.087*	X	0.121*	X
T13: Spare Oil Prod Cap	X	X	X	X	-0.025*	X
T17: Heavy Crude	X	X	-0.022*	0.019	-0.012*	0.029*
Energy policy signal						
T16: Energy Investments	X	X	X	0.045*	X	0.049*
T36: Energy Policy	X	X	-0.030*	-0.014	X	-0.005
Others signal						
T12: Oil Demand	0.042*	X	0.119*	X	0.059*	X
T18: Oil and Gas Mkt	X	X	-0.023*	X	-0.034*	X
T28: Nigeria Crude Oil	X	X	X	X	-0.015*	0.001
Private signals						
Consensus 3 months (SD)	0.144*		0.078*		X	
Consensus 12 months (SD)	X		X		X	
Control variables						
Production Decision: Cut	0.144*		0.148*		0.153*	
Production Decision: Increase	X		0.024		X	
Extraordinary Meetings	X		X		X	
OPEC Cooperation	X		0.027		0.055*	
Adjusted R^2	43%		52%		49%	

Note: This table reports the estimated coefficients of the penalized bootstrap Lasso regressions for price volatility (in percentage points) during high and low private noise (3 months) organized as clusters (see communities in Figure 3). We only report statistically significant topics. ∗ indicates significance at the 5% level. "X" indicates zero value coefficients.

Table 3: Effectiveness of public and private signals on trading positions in high and low private noise (3 months)

Signals	**P/M/D/P/U**		**MM**		**SD**		**OR**	
Public signals	*High*	*Low*	*High*	*Low*	*High*	*Low*	*High*	*Low*
Supply adju. signal								
T5: Rebalancing Market	X	-8.242*	X	X	X	3.241	X	X
T9: Cooperation	7.365*	12.66*	7.318*	X	-28.46*	-8.082	11.69*	6.835
T10: OPEC Producers	5.975	X	X	X	-11.40*	X	6.777*	X
T20: Prod Adjustment/COVID	X	7.985*	X	X	-14.35*	-5.460	14.26*	3.385
T25: OPEC Prod. Adjust.	X	4.440	11.89*	9.747*	-10.93*	-16.53*	X	X
T29: Minister Energy Negotiations	-3.929*	-4.425	-2.745	-1.953	7.856*	13.93*	X	X
T34: OPEC/NOPEC Prod. Part.	9.147*	0.009	14.94*	12.83*	-42.05*	-21.54*	14.61*	10.67*
OPEC internal. relations signal								
T15: OPEC-Russia Relations	8.385*	8.665*	X	X	X	X	X	X
T22: Intergovernmental Relations	X	7.541*	X	X	X	X	X	1.403
T33: Oil Industry	X	-7.608	X	12.49*	X	X	X	X
Price vol. & Mk stability signal								
T1: Extraodinary Meetings	-2.130	X	-6.692	X	19.86*	X	X	-2.705
T4: Oil Uncertainty/Volatility	X	X	-4.338	-10.96*	23.08*	28.23*	-6.300	-12.31*
T14: Oil Mkt Stability	-5.575*	-1.244	-1.518	X	X	6.114	4.003*	-7.737*
Prod.Ceiling & LT inv. signal								
T6: LT Strategy	0.216	-0.865	X	-2.511	X	14.35*	X	-6.892*
T21: Adequate Supply	X	X	-8.641*	-6.823*	23.07*	10.65*	-2.408	X
T37: Energy Security	X	1.138	X	-10.07*	X	X	X	X
T38: Petroleum Industry	-4.971*	X	-0.850	-7.7*	12.83*	17.50*	-0.143	-6.332*
T39: Production Ceiling	X	X	-5.010	X	10.95	X	-8.927*	X
T40: Global Prod. Capacity	X	X	X	-6.867*	X	16.54*	X	-8.244*
Oil shortage signal								
T11: Economic Growth	8.179*	10.48*	X	X	X	X	X	X
T13: Spare Oil Prod Cap	X	-1.794	X	-5.619*	X	5.684	X	X
T17: Heavy Crude	5.723	-2.270	X	X	X	16.49*	X	-9.255*
T19: Climate Change	-2.223*	-1.281	-2.337	X	5.573	4.601	-1.663	-1.360
Energy policy signal								
T36: Energy Policy	-7.537*	X	-5.375	-2.985	30.70*	16.86*	-8.123*	-10.84*
Others signal								
T7: Physical/Financial Interaction	9.600*	3.901	X	X	X	X	X	-0.086
T23: Kuwait Cooperation	2.310	-5.756	9.113*	6.172*	-9.355*	X	X	X
T30: Energy Outlook	X	3.105	2.181	12.49*	X	-18.19*	X	X
T35: World Economy	-7.388*	-2.481	X	-0.252	13.85*	17.76*	-6.146*	-9.704*
Private signals								
Consensus 3 months (SD)	28.97*		-6.949		X		X	
Consensus 12 months (SD)	X		-27.70*		49.81*		-48.29*	
Control variables								
Production Decision: Cut	-3.185*		X		X		X	
Production Decision: Increase	X		X		X		X	
Extraordinary Meetings	X		X		X		X	
OPEC Cooperation	-47.11*		X		53.26*		-9.263*	
Adjusted R^2	67%		49%		59%		67%	

Note: This table reports the estimated coefficients of the penalized bootstrap Lasso regressions for trading positions during high and low private noise (3 months) organized as clusters (see communities in Figure 3). We only report statistically significant topics. $*$ indicates significance at the 5% level. "X" indicates zero value coefficients. To save space, coefficients of trading positions are divided by 1000.

5.1.2 Robustness to different specifications

To assess the robustness of our findings, we estimate various alternative specifications: (i) including additional macroeconomic variables,[48] (ii) excluding the 2007-2008 boom-bust period, (iii) distinguishing between high and low OPEC spare capacity, and (iv) without normalizing dummy variables. We also investigate the effectiveness of public and private signals in high and low 12-month private noise. Finally, instead of considering the volatility on the announcement day, we computed the change in volatility with respect to the previous day.

The results from these alternative specifications are similar to those reported in the paper, illustrating the robustness of our findings. As an example, Table 11 in Appendix E reports the estimation results corresponding to the case where the ADS index has been replaced by Kilian (2009)'s index of global real economic activity (denoted as "dry cargo" in the table). As shown, they are identical to those we previously obtained.[49]

5.1.3 Selective inference problem

A challenging task when estimating high-dimensional statistics is to make inference accounting for uncertainty and hypothesis testing. This statistical problem is known as "selective inference" and raises concerns about the effects of variables' selection on inference.[50] Several methods exist to correct the problem.[51] In the core of the paper, we report the results from the residual bootstrap Lasso regression proposed by Chatterjee & Lahiri (2013). The bootstrap procedure allows us to measure parameters uncertainty in terms of confidence intervals.

To check the robustness of our results, we use numerous frequentist and Bayesian methods that proved to be efficient in high-dimensional setting. Tables 12 to 14 in Appendix E report the results from the following methods for Equation (7): (i) bootstrap Lasso + Partial Ridge (Liu et al. (2020), Table 12), (ii) bootstrap de-sparsed Lasso (Zhang & Zhang (2014), Table 13), and (iii) Bayesian Lasso (Park & Casella (2008), Table 14).[52] Our main conclusions remain valid regardless of the estimation method adopted.

[48] Specifically, we estimate three specifications: (i) a specification including the eight variables mentioned in Section 2.1, (ii) the same model in which the ADS index is replaced by the index of global real economic activity proposed by Kilian (2009), and (iii) a specification in which only the index of global real economic activity is included.

[49] To save space, we do not report all the results related to our robustness checks to different specifications, but they are available upon request from the authors.

[50] See, e.g., Taylor & Tibshirani (2015) for a discussion.

[51] See Dezeure et al. (2015) for a review of the most common existing methods.

[52] We also performed bootstrap Lasso + OLS (Liu & Yu (2013)), and multi sample-splitting methods. Results are robust and available upon request from the authors. Results for Equation (8) also corroborate our main conclusions and are available upon request.

5.2 Placebo test

As with all regression settings, an important identification assumption we make is that the responses of price volatility and trading positions we observe are the consequence of OPEC communications rather than the result of intrinsic dynamics in the oil market. To test the relevance of our identification structure, we conduct a placebo test during days with no OPEC announcements. We construct placebo samples by suppressing OPEC announcements days. From the placebo samples, we build control groups by sampling out with replacement 262 observations from futures price volatility over each maturity, and 208 observations from trading positions for each category of traders.[53]

For each control group, we repeat the estimation procedure of Equations (7) and (8).[54] Our results unequivocally indicate that all coefficients capturing OPEC announcements are zero (this is due to the use of LASSO penalized regressions). This finding shows that OPEC announcements are relevant when they occur and provide support to our identification strategy.

6 Conclusions

In this paper we are interested in analyzing the content of OPEC communications and whether it provides valuable information to the crude oil market. Starting from the Morris & Shin (2002) framework, we derive an empirical strategy which assumes that fundamental factors related to supply, demand, and speculative activity drive OPEC's public signal. Both public and private signals affect the crude oil market. Our results suggest that OPEC narratives cover a wide range of topics that are indeed linked to the fundamental factors we consider. We also find that OPEC narratives are relevant in the sense that they reduce crude oil price volatility and prompt market participants to rebalance their positions.

Our results stimulate further research. It would be interesting to know which market participants have the largest and fastest reactions to OPEC announcements. However, we recognize that data limitations pose an obstacle. In fact, to perform this analysis we would need access to confidential, detailed market participant positions. It would also be important to understand how our results in the crude oil market are transmitted to other markets. Crude oil is extremely important for both the real economy and financial markets. Understanding possible contagion mechanisms will help identify interconnectedness effects. Finally, our findings show that climate change is an important topic in OPEC communication. Studying how climate-related risks influence OPEC narratives and, in turn, the crude oil market, is critical to policymakers, market participants, and the general public.

[53] Recall, there are 262 and 208 OPEC announcement days when considering volatility (sample period 2002-2021) and traders' positions (sample period 2006-2021), respectively.

[54] The random placebo procedure is used to avoid any subjectivity in the choice of the pre-OPEC announcement days. As a robustness check, we also perform the same analysis by manually selecting pre-OPEC announcement days. Results are robust and available upon request.

Table 4: Effectiveness of communication during crisis periods

Signals	1-month	6-month	12-month	P/M/D/P/U	MM	SD	OR
Public signals							
Supply adju. signal							
T9: Cooperation	X	0.112*	0.138*	9.716*	X	-44.80*	19.47*
T10: OPEC Producers	X	X	-0.025	3.067*	X	0.480	X
T20: Prod Adjustment/COVID	X	0.227*	0.188*	X	X	-11.98	25.49*
T25: OPEC Prod. Adjust.	X	X	X	X	42.11*	-35.98*	3.941
T29: Minister Energy Negotiations	X	-0.030	-0.017	X	-3.534	37.09*	-7.907*
T34: OPEC/NOPEC Prod. Part.	X	X	X	X	X	X	-0.281
OPEC internal. relations signal							
T8: Iraq-Saudi Relations	X	0.026	0.068	X	10.83	-21.94	22.14*
Price vol. & Mk stability signal							
T1: Extraodinary Meetings	X	X	X	X	-4.319	69.31*	-19.02*
T4: Oil Uncertainty/Volatility	X	X	-0.070	-5.394*	-4.473	54.61*	-13.69*
T14: Oil Mkt Stability	0.567*	0.565*	0.306*	5.930*	X	-13.85	28.66*
Prod. Ceiling & LT inv. signal							
T6: LT Strategy	X	-0.046	X	-3.944	-1.480	25.41*	-4.942
T21: Adequate Supply	X	X	0.046*	X	X	X	X
T37: Energy Security	X	X	0.049*	X	-3.326	80.21*	-27.75*
T38: Petroleum Industry	X	-0.065*	-0.043*	X	-8.384*	31.32*	-6.688
T40: Global Prod. Capacity	X	-0.047	-0.126*	X	-6.035	33.17	-21.66*
Oil shortage signal							
T11: Economic Growth	X	X	X	X	X	-15.72*	X
Energy policy signal							
T36: Energy Policy	X	-0.083*	-0.101*	X	-8.765*	62.69*	-21.61*
Others signal							
T7: Physical/Financial Interaction	X	X	-0.0551*	X	X	-0.683	0.672
T18: Oil and Gas Mkt	X	X	X	X	X	X	5.326*
T23: Kuwait Cooperation	X	X	X	X	X	-7.947*	X
T28: Nigeria Crude Oil	X	X	X	X	1.495	-32.88*	2.300
T30: Energy Outlook	X	X	X	X	12.82*	-17.92*	X
T35: World Economy	X	X	X	X	12.66*	-24.38*	2.518
Private signals							
Consensus 3 months (SD)	X	X	X	1.624	-0.011	X	X
Consensus 12 months (SD)	X	X	X	X	-39.20*	53.47*	-28.17
Control variables							
Production Decision: Cut	X	X	0.006	X	X	X	-4.180*
Production Decision: Increase	X	X	-0.015	X	X	9.224	-9.725
Extraordinary Meetings	X	X	X	X	X	X	X
OPEC Cooperation	X	X	0.005	X	X	X	X
Adjusted R^2	44%	65%	68%	35%	74%	88%	88%

Note: This table reports the estimated coefficients of the penalized bootstrap Lasso regressions for both price volatility (in percentage points) and trading positions organized as clusters (see communities in Figure 3) during crisis periods. We only report statistically significant topics. $*$ indicates significance at the 5% level. "X" indicates zero value coefficients. To save space, coefficients of trading positions are divided by 1000.

References

Almoguera, P., Douglas, C. & Herrera, A. (2011), 'Testing for the cartel in OPEC: non-cooperative collusion or just non-cooperative?', *Oxford Review of Economic Policy*

27(1), 144–168.

Amato, J. D., Morris, S. & Shin, H. S. (2002), 'Communication and monetary policy', *Oxford Review of Economic Policy* **18**(4), 495–503.

Aruoba, S., Diebold, F. & Scotti, C. (2009), 'Real-time measurement of business conditions', *Journal of Business & Economic Statistics* **27**(4), 417–427.

Baumeister, C. & Kilian, L. (2016), 'Forty years of oil price fluctuations: Why the price of oil may still surprise us', *Journal of Economic Perspectives* **30**(1), 139–60.

Bischof, J. & Airoldi, E. M. (2012), Summarizing topical content with word frequency and exclusivity, *in* 'Proceedings of the 29th International Conference on Machine Learning (ICML-12)', pp. 201–208.

Blei, D. & Lafferty, J. (2009), 'Topic models', *Text mining: classification, clustering, and applications* **10**(71), 34.

Blei, D. M. (2012), 'Probabilistic topic models', *Communications of the ACM* **55**(4), 77–84.

Blei, D. M. & Lafferty, J. D. (2006), Dynamic topic models, *in* 'Proceedings of the 23rd international conference on Machine Learning', pp. 113–120.

Blei, D. M. & Lafferty, J. D. (2007), 'A correlated topic model of science', *The Annals of Applied Statistics* **1**(1), 17–35.

Blei, D., Ng, A. & Jordan, M. (2003), 'Latent dirichlet allocation', *Journal of Machine Learning Research* **3**(Jan), 993–1022.

Blondel, V. D., Guillaume, J.-L., Lambiotte, R. & Lefebvre, E. (2008), 'Fast unfolding of communities in large networks', *Journal of Statistical Mechanics: Theory and Experiment* **2008**(10), P10008.

Brémond, V., Hache, E. & Mignon, V. (2012), 'Does OPEC still exist as a cartel? an empirical investigation', *Energy Economics* **34**(1), 125–131.

Brunetti, C., Büyükşahin, B. & Harris, J. H. (2016), 'Speculators, prices, and market volatility', *Journal of Financial and Quantitative Analysis* **51**(5), 1545–1574.

Brunetti, C., Buyuksahin, B., Robe, M. & Soneson, K. (2013), 'OPEC "fair" price pronouncements and the market price of crude oil', *The Energy Journal* **34**, 79–108.

Brunetti, C. & Lildholdt, P. M. (2007), 'Time series modeling of daily log-price ranges for CHF/USD and USD/GBP', *The Journal of Derivatives* **15**(2), 39–59.

Calomiris, C. W. & Mamaysky, H. (2019), 'How news and its context drive risk and returns around the world', *Journal of Financial Economics* **133**(2), 299–336.

Chang, J., Gerrish, S., Wang, C., Boyd-Graber, J. & Blei, D. (2009), 'Reading tea leaves: How humans interpret topic models', *Advances in neural information processing systems* **22**, 288–296.

Chatterjee, A. & Lahiri, S. N. (2013), 'Rates of convergence of the adaptive LASSO estimators to the oracle distribution and higher order refinements by the bootstrap', *The Annals of Statistics* **41**(3), 1232–1259.

Chinco, A., Clark-Joseph, A. D. & Ye, M. (2019), 'Sparse signals in the cross-section of returns', *The Journal of Finance* **74**(1), 449–492.

Coenen, G., Ehrmann, M., Gaballo, G., Hoffmann, P., Nakov, A., Nardelli, S., Persson, E. & Strasser, G. (2017), Communication of monetary policy in unconventional times, Working Paper Series 2080, European Central Bank.
URL: *https://ideas.repec.org/p/ecb/ecbwps/20172080.html*

Demirer, R. & Kutan, A. M. (2010), 'The behavior of crude oil spot and futures prices around OPEC and SPR announcements: An event study perspective', *Energy Economics* **32**(6), 1467–1476.

Dezeure, R., Bühlmann, P., Meier, L. & Meinshausen, N. (2015), 'High-dimensional inference: confidence intervals, p-values and r-software hdi', *Statistical Science* pp. 533–558.

Eggertsson, G. B. & Woodford, M. (2003), 'Optimal monetary policy in a liquidity trap'.

Ehrmann, M. & Fratzscher, M. (2007), Social value of public information: testing the limits to transparency, Working Paper Series 821, European Central Bank.
URL: *https://ideas.repec.org/p/ecb/ecbwps/2007821.html*

Epps, T. & Epps, M. (1976), 'The stochastic dependence of security price changes and transaction volumes: Implications for the mixture-of-distributions hypothesis', *Econometrica* **44**(2), 305–321.

Evans, C., Fisher, J., Justiniano, A. & Campbell, J. (2012), 'Macroeconomic effects of FOMC forward guidance', *Brookings Papers on Economic Activity* .

Farrell, J. & Rabin, M. (1996), 'Cheap talk', *Journal of Economic Perspectives* **10**(3), 103–118.

Fattouh, B. (2007), *OPEC pricing power: The need for a new perspective*, Oxford Institute for Energy Studies.

Fattouh, B. & Mahadeva, L. (2013), 'OPEC: what difference has it made?', *Annu. Rev. Resour. Econ.* **5**(1), 427–443.

Freyberger, J., Neuhierl, A. & Weber, M. (2020), 'Dissecting characteristics nonparametrically', *The Review of Financial Studies* **33**(5), 2326–2377.

Gallant, R., Rossi, P. & Tauchen, G. (1992), 'Stock prices and volume', *The Review of Financial Studies* **5**(2), 199–242.

Griffiths, T. & Steyvers, M. (2004), 'Finding scientific topics', *Proceedings of the National Academy of Sciences* **101**(suppl 1), 5228–5235.

Gu, S., Kelly, B. & Xiu, D. (2020), 'Empirical asset pricing via machine learning', *The Review of Financial Studies* **33**(5), 2223–2273.

Guidi, M. G., Russell, A. & Tarbert, H. (2006), 'The effect of OPEC policy decisions on oil and stock prices', *OPEC Energy Review* **30**(1), 1–18.

Hansen, S. & McMahon, M. (2016), 'Shocking language: Understanding the macroeconomic effects of central bank communication', *Journal of International Economics* **99**, S114–S133.

Hansen, S., McMahon, M. & Prat, A. (2018), 'Transparency and deliberation within the FOMC: a computational linguistics approach', *The Quarterly Journal of Economics* **133**(2), 801–870.

Hermann, M. & Fratzscher, M. (2007a), 'Communication and decision-making by central bank committees: Different strategies, same effectiveness', *Journal of Money, Credit and Banking* **39**(2-3), 509–541.

Hermann, M. & Fratzscher, M. (2007b), 'Transparency, disclosure and the Federal Reserve', *International Journal of Central Banking* **3**(1), 179–225.

Hubert, P. & Labondance, F. (2018), 'The effect of ECB forward guidance on the term structure of interest rates', *International Journal of Central Banking* **56**.

Känzig, D. (2021), 'The macroeconomic effects of oil supply news: Evidence from OPEC announcements', *American Economic Review* **111**(4), 1092–1125.

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), 1053–69.

Kilian, L. (2023), How to Construct Monthly VAR Proxies Based on Daily Futures Market Surprises, Working Papers 2310, Federal Reserve Bank of Dallas.

Kilian, L. & Murphy, D. P. (2014), 'The role of inventories and speculative trading in the global market for crude oil', *Journal of Applied Econometrics* **29**(3), 454–478.

Kozak, S., Nagel, S. & Santosh, S. (2020), 'Shrinking the cross-section', *Journal of Financial Economics* **135**(2), 271–292.

Larsen, V. H. & Thorsrud, L. A. (2019), 'The value of news for economic developments', *Journal of Econometrics* **210**(1), 203–218.

Lin, S. & Tamvakis, M. (2010), 'OPEC announcements and their effects on crude oil prices', *Energy Policy* **38**(2), 1010–1016.

Liu, H., Xu, X. & Li, J. J. (2020), 'A bootstrap lasso+ partial ridge method to construct confidence intervals for parameters in high-dimensional sparse linear models', *Statistica Sinica* **30**(3), 1333–1355.

Liu, H. & Yu, B. (2013), 'Asymptotic properties of lasso+ mLS and lasso+ ridge in sparse high-dimensional linear regression', *Electronic Journal of Statistics* **7**, 3124–3169.

Loutia, A., Mellios, C. & Andriosopoulos, K. (2016), 'Do OPEC announcements influence oil prices?', *Energy Policy* **90**, 262–272.

Meilă, M. (2003), Comparing clusterings by the variation of information, *in* 'Learning theory and kernel machines', Springer, pp. 173–187.

Mimno, D., Wallach, H., Talley, E., Leenders, M. & McCallum, A. (2011), Optimizing semantic coherence in topic models, *in* 'Proceedings of the 2011 conference on empirical methods in natural language processing', pp. 262–272.

Morris, S. & Shin, H. S. (2002), 'Social value of public information', *American Economic Review* **92**(5), 1521–1534.

Morris, S. & Shin, H. S. (2005), 'Central bank transparency and the signal value of prices', *Brookings Papers on Economic Activity* **2005**(2), 1–66.

Morris, S., Shin, H. S. & Tong, H. (2006), 'Social value of public information: Morris and shin (2002) is actually pro-transparency, not con: Reply', *American Economic Review* **96**(1), 453–455.

Park, T. & Casella, G. (2008), 'The bayesian lasso', *Journal of the American Statistical Association* **103**(482), 681–686.

Pescatori, M. A. & Nazer, Y. F. (2022), OPEC and the Oil Market, IMF Working Papers 2022/183, International Monetary Fund.

Pons, P. & Latapy, M. (2006), Computing communities in large networks using random walks, *in* 'J. Graph Algorithms Appl', Citeseer.

Quint, D. & Venditti, F. (2020), 'The influence of OPEC+ on oil prices: a quantitative assessment', *ECB Working Paper 2467* .

Raghavan, U. N., Albert, R. & Kumara, S. (2007), 'Near linear time algorithm to detect community structures in large-scale networks', *Physical Review E* **76**(3), 036106.

Roberts, M. E., Stewart, B. M. & Tingley, D. (2016), 'Navigating the local modes of big data', *Computational Social Science* **51**.

Roberts, M. E., Stewart, B. M. & Tingley, D. (2019), 'Stm: An R package for structural topic models', *Journal of Statistical Software* **91**(1), 1–40.

Roberts, M. E., Stewart, B. M., Tingley, D., Airoldi, E. M. et al. (2013), The structural topic model and applied social science, *in* 'Advances in neural information processing systems workshop on topic models: computation, application, and evaluation', Vol. 4, Harrahs and Harveys, Lake Tahoe, pp. 1–20.

Roberts, M. E., Stewart, B. M., Tingley, D., Lucas, C., Leder-Luis, J., Gadarian, S. K., Albertson, B. & Rand, D. G. (2014), 'Structural topic models for open-ended survey responses', *American Journal of Political Science* **58**(4), 1064–1082.

Rosvall, M. & Bergstrom, C. T. (2007), 'Maps of information flow reveal community structure in complex networks', *arXiv preprint physics.soc-ph/0707.0609* .

Sanders, D. & Irwin, S. (2013), 'Measuring index investment in commodity futures markets', *The Energy Journal* **34**(3), 105–127.

Scotti, C. (2016), 'Surprise and uncertainty indexes: Real-time aggregation of real-activity macro-surprises', *Journal of Monetary Economics* **82**, 1–19.

Svensson, L. E. (2006), 'Social value of public information: Comment: Morris and shin (2002) is actually pro-transparency, not con', *American Economic Review* **96**(1), 448–452.

Taddy, M. (2012), On estimation and selection for topic models, *in* 'Artificial Intelligence and Statistics', PMLR, pp. 1184–1193.

Tauchen, G. & Pitts, M. (1983), 'The price variability-volume relationship on speculative markets', *Econometrica* **51**(2), 485–505.

Taylor, J. & Tibshirani, R. J. (2015), 'Statistical learning and selective inference', *Proceedings of the National Academy of Sciences* **112**(25), 7629–7634.

Tibshirani, R. (1997), 'The lasso method for variable selection in the Cox model', *Statistics in Medicine* **16**(4), 385–395.

Wallach, H. M., Murray, I., Salakhutdinov, R. & Mimno, D. (2009), Evaluation methods for topic models, *in* 'Proceedings of the 26th annual international conference on machine learning', pp. 1105–1112.

Wirl, F. & Kujundzic, A. (2004), 'The impact of OPEC conference outcomes on world oil prices 1984-2001', *The Energy Journal* **25**(1).

Zhang, C.-H. & Zhang, S. S. (2014), 'Confidence intervals for low dimensional parameters in high dimensional linear models', *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **76**(1), 217–242.

Appendix

A Pre-processing, and model selection

As discussed in Section 2.3, dimensionality reduction is a key first step when using computational linguistic models since text datasets are often both large and scarce. Table 5 reports the pre-processing steps used in the paper with the evolution of total words. Our corpus starts with 12586 words and ends with 2573 meaningful words after preprocessing.

Table 5: Data dimensionality reduction of each preprocessing step

	Raw Text	Remove Stopwords & Given Names	Remove Numbers & Punctuations	Remove words $<$ 3 length + Stemming Algorithm
Total words	12586	11883	7400	5027

Notes: The table reports the evolution of total words through pre-processing. The stemming algorithm is the Porter stemmer implemented in R using 'tm' package.

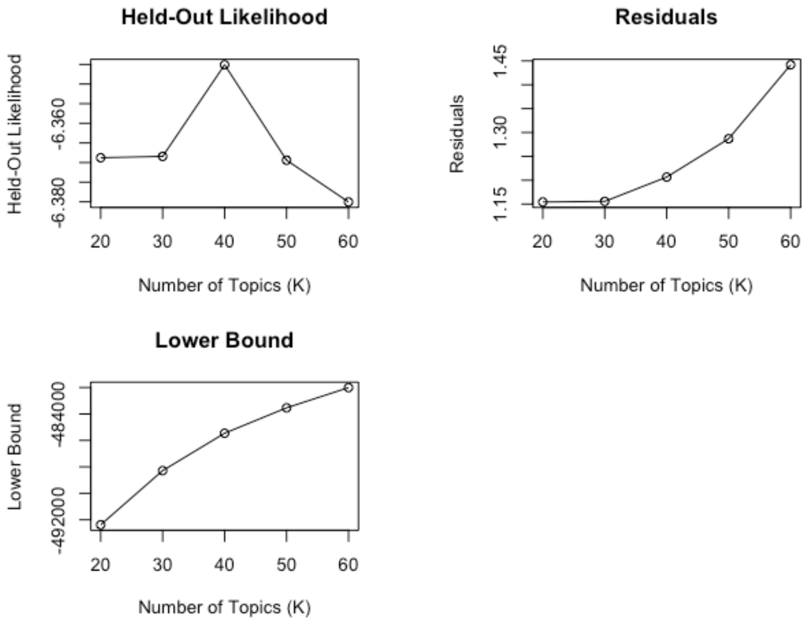
Another important element in estimating topic models is the number K of topics. We use several methods to help choosing the number of topics. Figure 11 reports our considered evaluation measures for Topics 20 to 60, such as the held-out likelihood (Wallach et al. (2009)), the residual checks (Taddy (2012)), and the lower bound.[55] One needs to find the right trade-off for all measures, namely the number of topics for which each considered criteria is reasonably good. Based on our diagnostic, we select $K = 40$. Another selection criteria used in the literature is the semantic coherence developed by Mimno et al. (2011).[56] As noted by Roberts et al. (2014), semantic coherence alone is relatively easy to achieve by having only a couple of topics which all are dominated by the most common words. We therefore follow Roberts et al. (2014) and report in Figure 12 a combination of semantic coherence and exclusivity of words to topics.[57] The coherence-exclusivity trade-off confirms our choice of $K = 40$.

[55] For a discussion on each measure, see Roberts et al. (2019).

[56] Semantic coherence is related to pointwise mutual information and is maximized when the most probable words in a given topic frequently co-occur together.

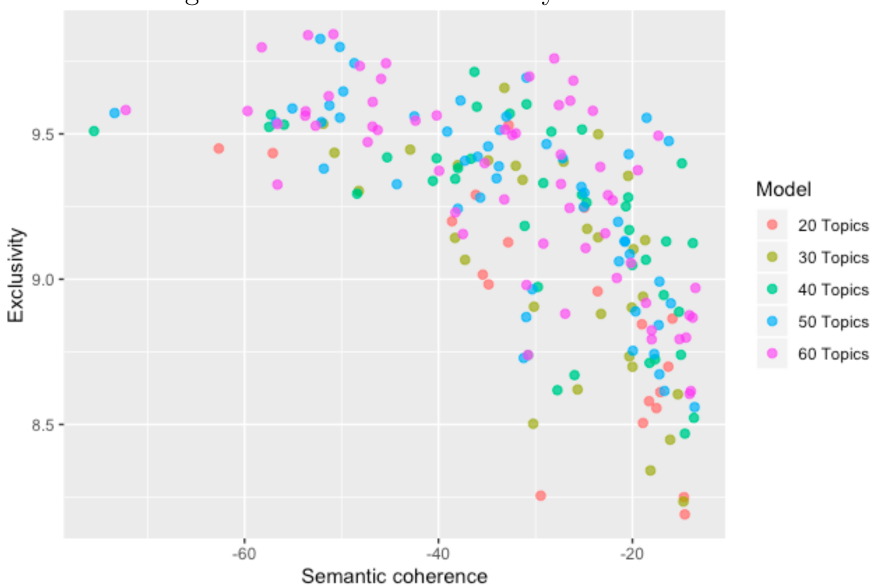
[57] In our case, exclusivity is measured by FREX metric (see Bischof & Airoldi (2012)). Section B in Appendix discusses in more details the FREX measure.

Figure 11: Diagnostic values by number of topics



Note: This figure reports different measures of topic selection for several topics values (from 20 to 60). Both held-out likelihood and lower bound have to be maximized, while residual diagnostic need to be minimized.

Figure 12: Semantic exclusivity vs coherence



B Topic labeling

This section briefly presents the two approaches we use for topic labeling. Recall that labels play no role in the analysis but provide a convenient way to discuss our results. For each of the 40 topics, we first use the FREX metric defined as the weighted harmonic mean of the word's rank in terms of exclusivity and frequency:

$$FREX_{k,\nu} = \left(\frac{\omega}{ECDF\left(\beta_{k,\nu}/\sum_{j=1}^{K}\beta_{j,\nu}\right)} + \frac{1-\omega}{ECDF\left(\beta_{k,\nu}\right)} \right)^{-1}$$

where $ECDF$ is the frequency score given the empirical CDF of the word in its topic distribution. ω is the weight sets to 0.7 (to favor exclusivity). Exclusivity is calculated by normalizing the β matrix (i.e., the conditional probability of topics given the word). Words with high value are those where most of the mass for these words is assigned to the given topic. Together with FREX, we also use the most-probable bigrams.[58] Both metrics are reported for each topic in Tables 6 to 8.

[58] A bigram is an association of two words.

Table 6: Estimated topics and labeling (Topic 1 to 15)

Topics	Label	Top 10 terms
Topic 1	Extraordinary meetings	meet, market, opec, organ, petroleum, current, countri, extraordinari, republ, suppli
Topic 2	Basket price	basket, refer, wti, cut, crude, barrel, quarter, averag, russia, month
Topic 3	Oil shortage	howev, quarter, extraordinari, wish, like, ceil, level, can, purpos, deepest
Topic 4	Oil uncertainty/volatility	volatil, specul, fundament, crude, geopolit, oilpric, price, increas, day, comfort
Topic 5	Rebalancing market	algier, accord, agreement, committe, algeria, forward, rebalanc, overhang, reactiv, high-level
Topic 6	Long-term strategy	strategi, long-term, object, futur, consist, identifi, adopt, multilater, role, technolog
Topic 7	Physical/Financial interaction	physic, workshop, financi, ief, interact, evolv, iea, regul, three, event
Topic 8	Iraq-Saudi relations	iraq, visit, iraqi, said, algier aramco, achiev, prime, venezuela, extens
Topic 9	Cooperation	committe, declar, nopec, technic, monitor, adjust, voluntari, joint, particip, return
Topic 10	OPEC producers	join, sovereign, declar, peopl, join, right, organ, cooper, withdraw, nation
Topic 11	Economic growth	growth, barrel, economic growth, averag, project, year, forecast, like, oecd, balanc
Topic 12	Oil demand	like, locat, recoveri, sign, posit, move, citi, district, general, libyan
Topic 13	Spare oil prod. capacity	capac, increas, rise, spare, downstream, avail, around, product, addit, crude oil
Topic 14	Oil market stability	countri, market, oilmarket, opec, oil, meet, stabil, member, world, global
Topic 15	OPEC-Russia relations	india, high-level, parti, russianf, long-term, opec-russia, dialogu, meet, senior, technic

Note: This table reports labels for Topics 1 to 15 based on both most probable bigrams (column “Label”) and top 10 FREX terms (column “Top 10 terms”).

Table 7: Estimated topics and labeling (Topic 16 to 30)

Topics	Label	Top 10 terms
Topic 16	Energy investments	market, price, oil, opec, suppli, invest, produc, consum, energi, oilpric
Topic 17	Heavy crude	composit, eleven, heavier, distil, orb, iran, trial, weight, temporarili, index
Topic 18	Oil and gas market	data, media, tool, avail, big, exercis, phase, project, uae, statist
Topic 19	Climate change	climat, chang, pari, convent, framework negoti, agreement, diversif, implement, sustain
Topic 20	Prod. adjustment / COVID	adjust, reaffirm, epidem, COVID, declar, compens, outbreak, particip, product, agre
Topic 21	Adequate supply	adequ, suppli, situat, level, price, band, close, light, qatar, consum
Topic 22	Intergovernmental Relations	gecf, area, data, experi, mutual, gas, exchang, two, common, sign
Topic 23	Kuwait cooperation	kuwait, declar, implement, prime, met, earlier, visit, cooper, role, congratul
Topic 24	Natural disaster	hurrican, condol, peopl, katrina, devast, caus, unit, sad, capac, govern
Topic 25	OPEC production adjustments	committe, compens, conform, month, particip, adjust, schedul, full, overal, rebalanc
Topic 26	Energy poverty	believ, prove, humankind, back, abl, statement, challenge, reflect, histori, togeth
Topic 27	OPEC-MENA/China relations	china, workshop, iea, region, opec-china, uncertainti, prospect, dialogu, demand, mena
Topic 28	Nigeria crude oil	nigeria, univers, oil, nation, serv, nigerian, former, petroleum, institut, gas
Topic 29	Minister energy negotiations	iran, qatar, algeria, current, consult, accord, restor, attend, negat, oilmarket
Topic 30	Energy Outlook	ief, compar, iea, energi, outlook, scenario, transit, iea, transpar, agenc

Note: This table reports labels for Topics 16 to 30 based on both most probable bigrams (column “Label”) and top 10 FREX terms (column “Top 10 terms”).

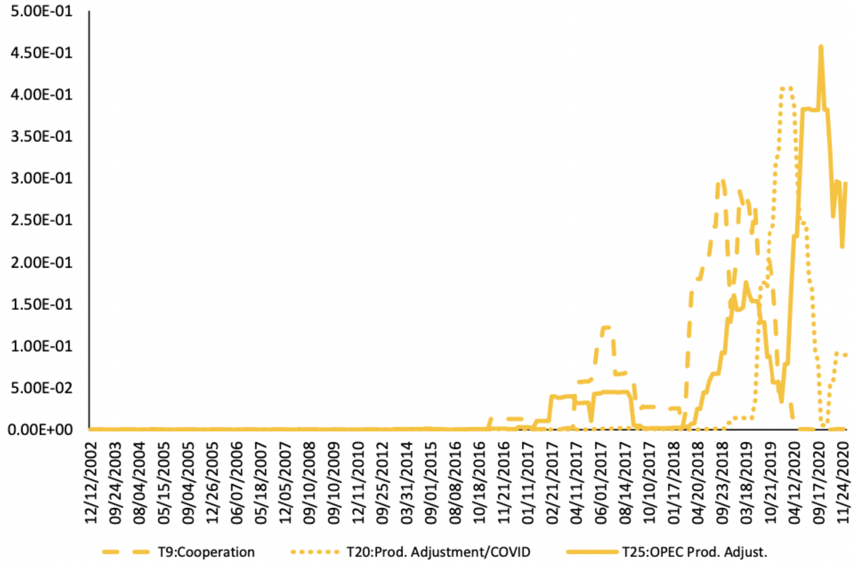
Table 8: Estimated topics and labeling (Topic 31 to 40)

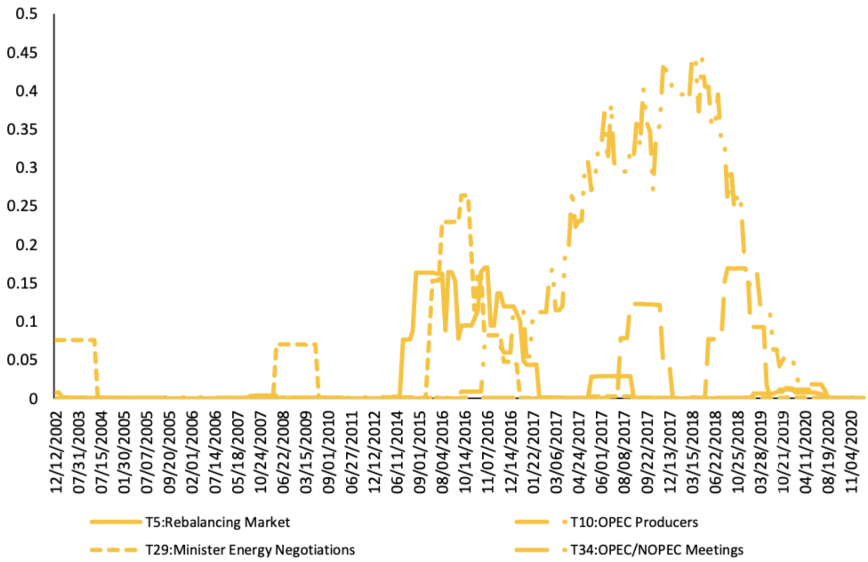
Topics	Label	Top 10 terms
Topic 31	OPEC/Asia dialogue	india, dialogu, iea, visit, pandem, research, energi, center, exchang, cooper
Topic 32	Gas & coal markets	coal, unit, gas, visit, center, commod, state, oil, outlook, imf
Topic 33	Oil industry	compani, ceo, india, sector, invest, offici, total, industri, spoke, eni
Topic 34	OPEC/Non-OPEC production participation	meet, declar, adjust, particip, cooper, joint, voluntari, opec-nopec, month, produc
Topic 35	World Economy	special, ceremoni, recess, exhibit, activ, govern, packag, golden, stamp, perform
Topic 36	Energy policy	energy, opec, technolog, polici, brussel, european, progress, dialogu, demand, fuelenergi,
Topic 37	Energy security	south, africa, osaka, shall, japan, cop, poverti, secur, kyoto, protocol
Topic 38	Petroleum industry	ministri, egypt, meet, let, oil-produc, shall, observ, come, scientif, ministri
Topic 39	Production ceiling	factor, reason, price, geopolit, pressur, stabil, specul, ceil, measur, tension
Topic 40	Global prod. capacity	nigeria, gas, reserv, doha, therefor, domin, price, noc, proud, polit

Note: This table reports labels for Topics 31 to 40 based on both most probable bigrams (column "Label") and top 10 FREX terms (column "Top 10 terms").

C Time evolution of OPEC signals

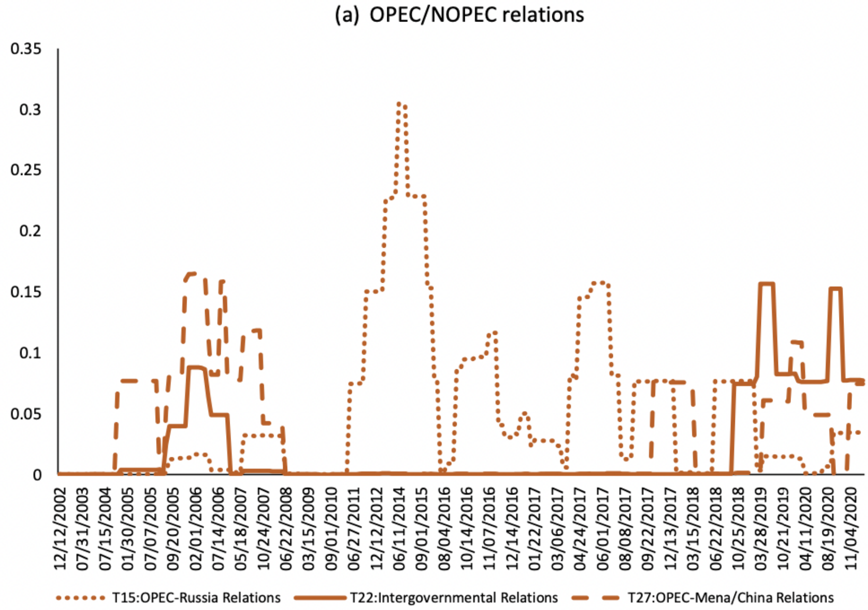
Figure 13: Supply adjustment signal evolution (orange community)

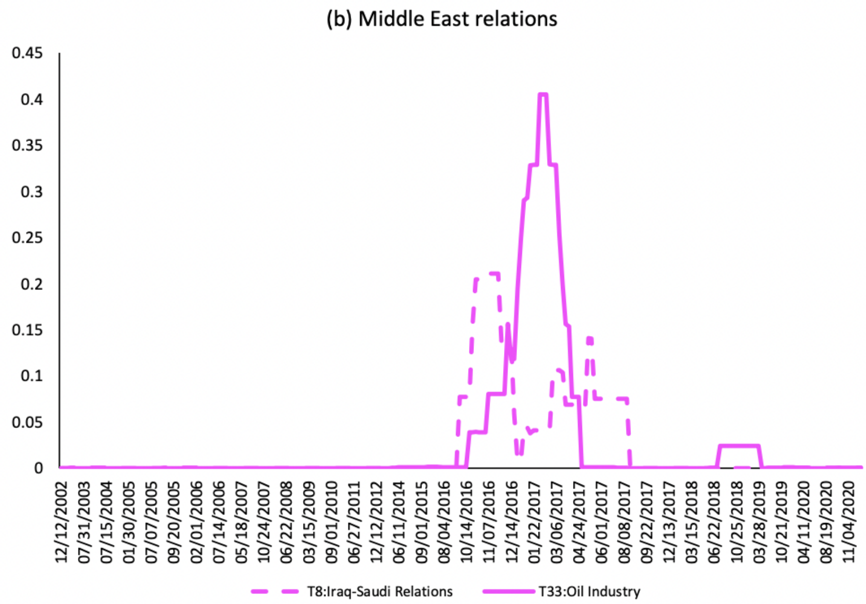




Note: This figure represents the topics probability over time in the orange community using a kernel smoothing transformation (Daniell method). The window size is 6 points which roughly corresponds to 6-months period.

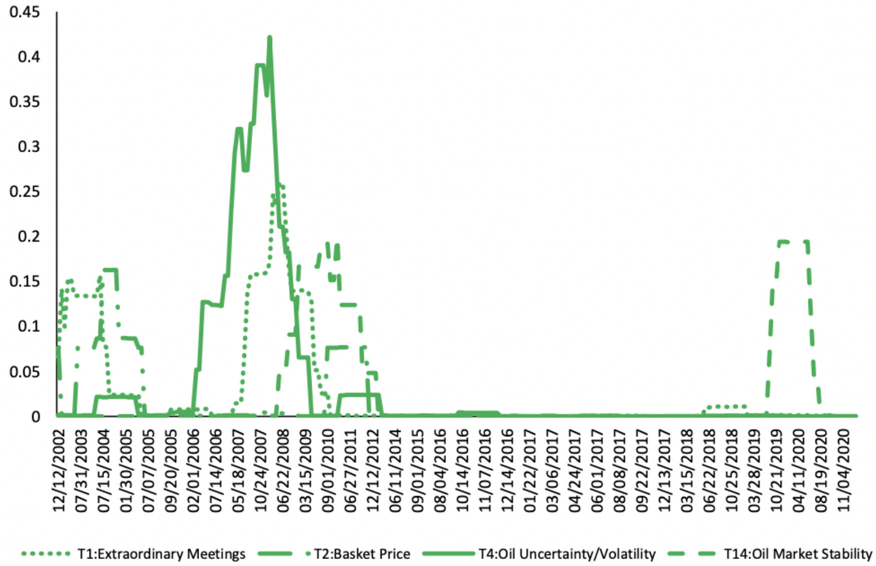
Figure 14: OPEC international relations signal (brown and purple communities)





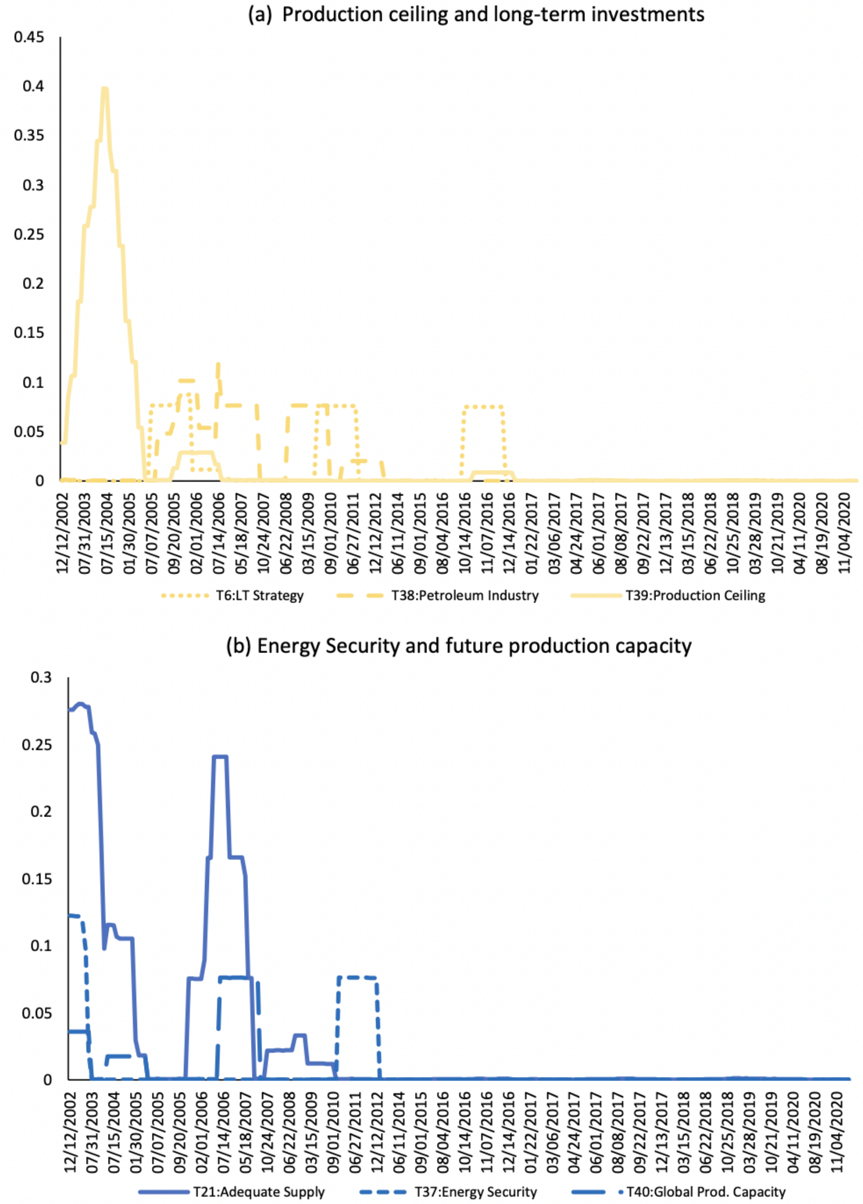
Note: This figure represents the topics probability over time in the brown (panel (a)) and purple (panel (b)) communities using a kernel smoothing transformation (Daniell method). The window size is 6 points which roughly corresponds to 6-months period.

Figure 15: Price volatility and market stability signal (green community)



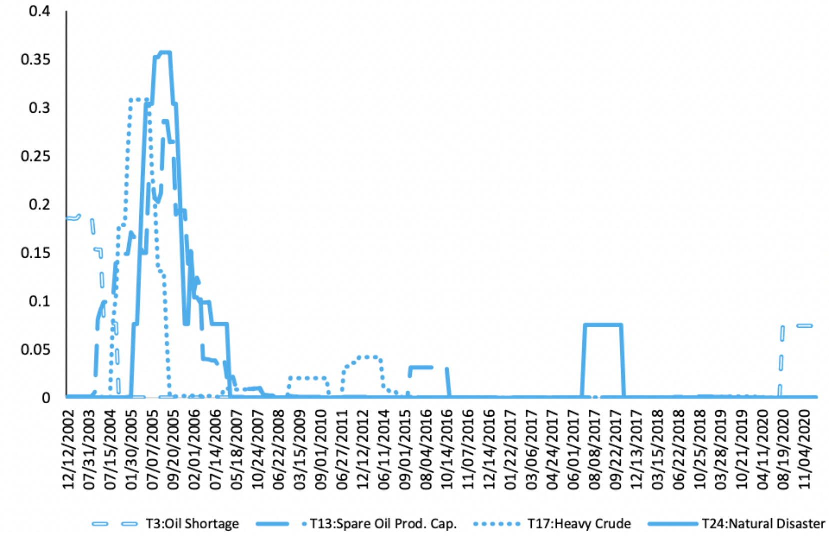
Note: This figure represents the topics probability over time in the green community using a kernel smoothing transformation (Daniell method). The window size is 6 points which roughly corresponds to 6-months period.

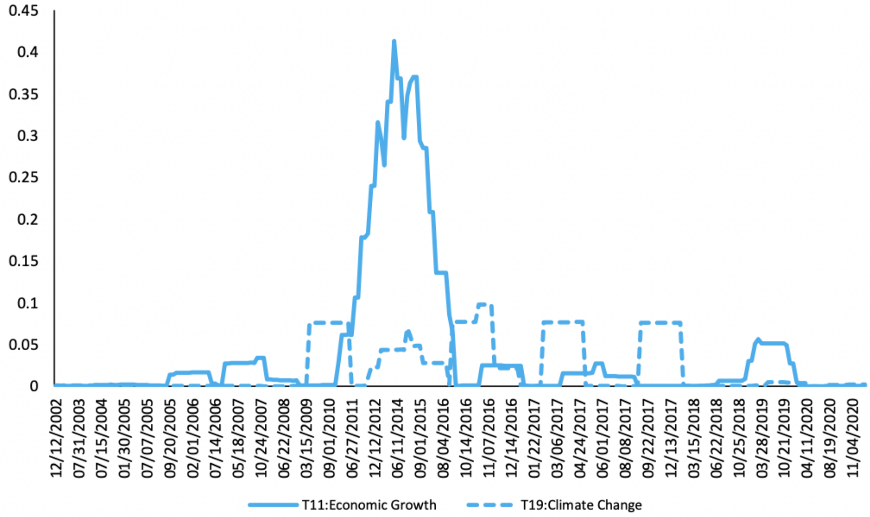
Figure 16: Long-term investment signal (yellow and darkblue communities)



Note: This figure represents the topics probability over time in the yellow (panel (a)) and dark blue (panel (b)) communities using a kernel smoothing transformation (Daniell method). The window size is 6 points which roughly corresponds to 6-months period.

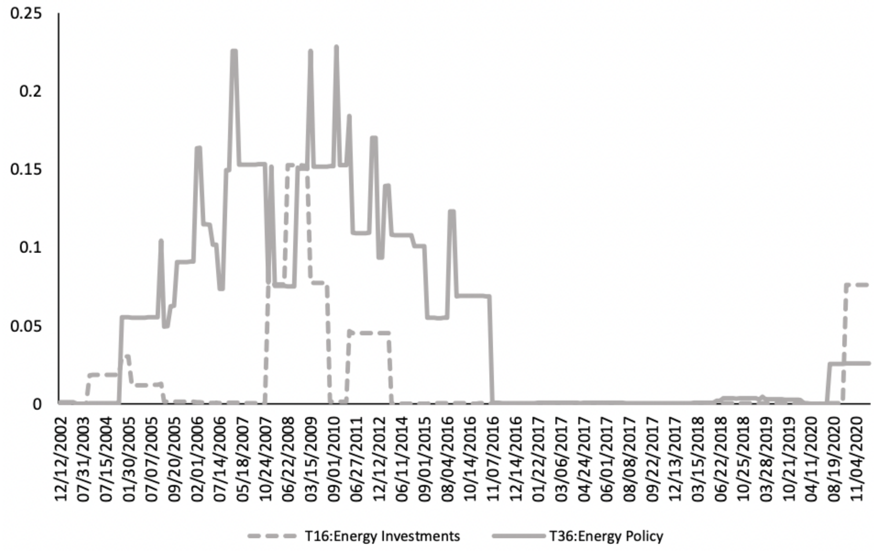
Figure 17: Oil shortage signal (light blue community)





Note: This figure represents the topics probability over time in the light blue community using a kernel smoothing transformation (Daniell method). The window size is 6 points which roughly corresponds to 6-months period.

Figure 18: Energy policy signal (grey community)



Note: This figure represents the topics probability over time in the grey community using a kernel smoothing transformation (Daniell method). The window size is 6 points which roughly corresponds to 6-months period.

D OPEC communication in recent crises

Table 9: Effectiveness of communication on price volatility in high and low private noise during crisis periods (3 months)

Signals	**1-month**		**6-month**		**12-month**	
Public signals	*High*	*Low*	*High*	*Low*	*High*	*Low*
Supply adju. signal						
T9: Cooperation	X	X	0.056	0.032	0.073*	0.090*
T10: OPEC Producers	X	X	X	X	-0.059*	0.045*
T20: Prod Adjustment/COVID	0.308*	X	0.184*	X	0.219*	0.061*
T25: OPEC Prod. Adjust.	X	X	0.061*	X	0.126*	X
T29: Minister Energy Negotiations	0.009	X	X	X	X	X
T34: OPEC/NOPEC Prod. Part.	0.161*	X	X	X	X	-0.009
OPEC internal. Relations signal						
T8 : Iraq-Saudi Relations	X	X	X	0.007	X	0.094*
T22: Inter-governmental Relations	X	X	0.001*	X	0.077*	X
Price vol. & Mk stability signal						
T2: Basket Price	X	X	X	-0.043*	X	-0.037*
T4: Oil Uncertainty/Volatiltiy	X	X	X	X	X	-0.037*
T14: Oil Mkt Stability	0.007*	X	0.411*	-0.003	0.290*	-0.053*
Prod.Ceiling & LT inv. signal						
T21: Adequate Supply	X	X	X	X	0.045*	-0.088*
T37: Energy Security	X	X	X	-0.020	X	X
T38: Petroleum Industry	X	X	X	-0.042*	0.011*	-0.017
T39: Production Ceiling	X	X	X	-0.017	0.022*	-0.006
T40: Global Prod. Capacity	X	X	X	X	-0.061*	X
Oil shortage signal						
T13: Spare Oil Prod Cap	X	X	X	X	-0.025	-0.014*
T17: Heavy Crude	X	X	X	X	-0.006	0.084*
Energy policy signal						
T16: Energy Investments	X	X	X	0.043	X	0.102*
Others signal						
T35: World Economy	X	X	X	-0.030*	-0.044*	X
Private signals						
Consensus 3 months (SD)	0.084		X		0.015	
Consensus 12 months (SD)	X		X		X	
Control variables						
Production Decision: Cut	0.037		0.028		0.068	
Production Decision: Increase	X		X		0.009	
Extraordinary Meetings	X		X		0.052	
OPEC Cooperation	X		X		X	
Adjusted R^2	64%		59%		75%	

Note: This table reports the estimated coefficients of the penalized bootstrap Lasso regressions for price volatility (in percentage points) during high and low private noise (3 months) organized as clusters (see communities in Figure 3) during crisis period. We only report statistically significant topics. $*$ indicates significance at the 5% level. "X" indicates zero value coefficients.

Table 10: Effectiveness of public and private signals on trading positions in high and low private noise during crisis periods (3 months)

Signals	P/M/D/P/U		MM		SD		OR	
Public signals	*High*	*Low*	***High***	***Low***	*High*	*Low*	*High*	*Low*
Supply adju. signal								
T9: Cooperation	12.59*	X	X	X	-36.32*	X	X	X
T10: OPEC Producers	4.034*	X	-1.234	X	X	X	X	X
T20: Prod Adjustment/COVID	X	X	X	X	-10.08	X	120.39*	X
T25: OPEC Prod. Adjust.	X	X	30.34*	11.28	-10.94	-4.359	X	X
T34: OPEC/NOPEC Prod. Part.	X	17.34*	X	X	X	-2.223	X	X
OPEC internal. relations signal								
T15: OPEC-Russia Relations	X	X	-6.003	X	15.89*	X	X	X
T22: Intergovernmental Relations	X	X	12.39*	X	X	X	X	X
T27: OPEC-MENA/China relations	X	X	20.62*	-23.83*	-4.556*	75.39*	X	-1.865
T33: Oil Industry	X	X	-22.61*	3.002	X	X	X	X
Price vol. & Mk stability signal								
T1: Extraordinary Meetings	-3.051*	X	-19.93*	X	45.49*	X	-48.65	X
T2 : Basket Price	X	-1.696	X	-16.37*	X	44.78*	X	X
T4: Oil Uncertainty/Volatility	-3.507	X	-19.29*	-18.27*	53.16*	61.24*	-72.68*	-175.8*
T14: Oil Mkt Stability	7.999*	-0.812	X	X	-19.02*	X	167.95*	X
Prod.Ceiling & LT inv. signal								
T6: LT Strategy	-2.485	X	-3.098	X	35.54*	X	X	X
T37: Energy Security	X	X	-3.771	-6.783	28.70*	6.064	X	X
T38: Petroleum Industry	-1.095	X	X	-12.05*	X	27.99*	X	-52.69*
T40: Global Prod. Capacity	X	3.326*	X	X	23.63*	X	X	X
Oil shortage signal								
T11: Economic Growth	X	X	2.109*	X	-7.310*	X	X	X
T19: Climate Change	X	X	0.634*	X	X	X	X	X
T24: Natural Disaster	16.61*	X	-0.175	-7.702*	5.997	X	X	X
Energy policy signal								
T16: Energy Investments	X	X	-2.503*	X	14.01*	X	X	X
T36: Energy Policy	X	-4.497	X	-35.80*	X	98.00*	X	-108.4*
Others signal								
T7: Physical/Financial Interaction	X	X	-10.91*	X	32.22*	X	X	X
T28: Nigeria Crude Oil	X	0.930	X	19.81*	X	-53.41*	X	X
T30: Energy Outlook	X	0.596	X	10.32*	X	-0.210	X	X
T35: World Economy	-9.737*	X	1.826	X	X	X	X	X
Private signals								
Consensus 3 months (SD)	2.356		-9.611*		X		-3.178	
Consensus 12 months (SD)	X		-45.58*		115.3*		-24.53*	
Control variables								
Production Decision: Cut	X		5.074*		X		X	
Production Decision: Increase	X		X		X		X	
Extraordinary Meetings	X		X		X		21.10	
OPEC Cooperation	X		X		X		-17.81	
Adjusted R^2	50%		88%		90%		78%	

Note: This table reports the estimated coefficients of the penalized bootstrap Lasso regressions for trading positions during high and low private noise (3 months) organized as clusters (see communities in Figure 3). We only report statistically significant topics. $*$ indicates significance at the 5% level. "X" indicates zero value coefficients. To save space, coefficients of trading positions are divided by 1000.

E Robustness checks

E.1 Macro specification

Table 11: Effectiveness of public and private signals (macro specification)

Signals	1-month	6-month	12-month	P/M/D/P/U	MM	SD	OR
Public signals							
Supply adju. signal							
T5: Rebalancing Market	X	0.029	X	-7.403*	X	X	X
T9: Cooperation	X	0.054*	X	15.83*	X	-32.72*	16.20*
T10: OPEC Producers	X	X	X	X	X	-9.743	7.371*
T20: Prod Adjustment/COVID	0.151*	X	X	X	X	-23.46*	16.00*
T25: OPEC Prod. Adjust.	X	X	X	X	18.35*	-26.95*	1.346
T29: Minister Energy Negotiations	0.090*	X	X	X	-5.415	23.42*	-1.999
T34: OPEC/NOPEC Prod. Part.	-0.002	X	X	X	21.54*	-43.75*	19.55*
OPEC internal. relations signal							
T15: OPEC-Russia Relations	X	X	X	12.62*	X	X	X
T33: Oil Industry	X	X	X	-1.755	14.77*	X	X
Price vol. & Mk stability signal							
T1: Extraodinary Meetings	X	X	X	-2.485	-9.894*	27.87*	-6.334
T4: Oil Uncertainty/Volatility	-0.040*	-0.042*	-0.049*	X	-15.17*	45.76*	-20.77*
T14: Oil Mkt Stability	0.345*	0.179*	0.209*	-7.131*	-2.202	5.662	X
Prod. Ceiling & LT inv. signal							
T6: LT Strategy	X	X	X	X	-2.162	16.23*	-8.490*
T21: Adequate Supply	X	X	-0.061*	X	-13.41*	29.08*	-11.30*
T38: Petroleum Industry	X	X	-0.031*	X	-9.307*	29.85*	-8.688*
T40: Global Prod. Capacity	-0.040*	-0.065*	-0.033*	X	-6.961*	18.49*	-10.05*
Oil shortage signal							
T11: Economic Growth	X	0.027	0.062*	14.71*	X	X	X
Energy policy signal							
T36: Energy Policy	-0.058*	-0.056*	-0.021*	X	-9.609*	42.66*	-19.07*
Others signal							
T7: Physical/Financial Interaction	-0.016	-0.023	-0.012	8.083*	X	X	X
T23: Kuwait Cooperation	X	X	X	-2.506	8.605*	X	X
T30: Energy Outlook	X	X	X	X	14.95*	-19.58*	X
T35: World Economy	X	X	X	-8.188*	-1.468	30.08*	-15.13*
Private signals							
Consensus 3 months (SD)	0.292*	0.126*	0.053*	29.07*	-5.832	X	15.66*
Consensus 12 months (SD)	X	X	X	X	-20.00*	38.43*	-53.31*
Control variables							
Production Decision: Cut	2.143*	0.150*	0.150*	X	X	X	X
Production Decision: Increase	X	0.014	X	X	X	X	-3.850
Extraordinary Meetings	X	X	X	X	X	X	X
OPEC Cooperation	X	0.049*	0.047	-48.16*	X	48.98*	-8.398*
Dry Cargo	1.051*	1.060*	1.047*	30.36*	X	48.70*	X
Adjusted R^2	37%	33%	33%	56%	44%	58%	68%

Note: This table reports the estimated coefficients of the penalized bootstrap Lasso for both price volatility (in percentage points) and trading positions organized as clusters (see communities in Figure 3). We only report statistically significant topics. $*$ indicates significance at the 5% level. "X" indicates zero value coefficients. To save space, coefficients of trading positions are divided by 1000. "Dry Cargo" is the index of global real economic activity developed by Kilian (2009).

E.2 Selective inference

Table 12: Bootstrap Lasso + Partial Ridge estimation

Signals	1-month	6-month	12-month	P/M/D/P/U	MM	SD	OR
Public signals							
Supply adju. signal							
T5: Rebalancing Market	0.037	0.032	0.028	-7.605*	X	7.818	X
T9: Cooperation	X	0.058*	0.038	16.13*	0.594	-32.23*	16.28*
T10: OPEC Producers	X	X	0.012	5.373	X	-8.317	7.484*
T20: Prod Adjustment/COVID	0.163*	0.046	0.039	3.994	X	-22.68*	15.89*
T25: OPEC Prod. Adjust.	X	X	0.003	2.603	19.14*	-26.65*	1.367
T29: Minister Energy Negotiations	0.106*	0.019	0.022	-7.069	-6.411	19.75*	-2.298
T34: OPEC/NOPEC Prod. Part.	-0.041	-0.016	X	4.466	22.17*	-43.71*	19.62*
OPEC internal. relations signal							
T15: OPEC-Russia Relations	X	X	X	12.887*	X	X	X
T33: Oil Industry	X	X	X	-8.148	15.57*	X	0.105
Price vol. & Mk stability signal							
T1: Extraodinary Meetings	X	-0.026	X	-2.769	-10.87*	23.66*	-6.676
T4: Oil Uncertainty/Volatility	-0.063*	-0.046*	-0.076*	X	-15.93*	41.95*	-21.02*
T14: Oil Mkt Stability	0.353*	0.182*	0.124*	-7.488*	-3.188	0.655	X
Prod. Ceiling & LT inv. signal							
T6: LT Strategy	X	X	-0.009	-1.257	-3.130	12.71*	-8.752*
T21: Adequate Supply	0.015	X	-0.066*	X	-14.39*	25.16*	-11.61*
T38: Petroleum Industry	0.0015	X	-0.038*	-3.936	-10.45*	25.56*	-8.912*
T40: Global Prod. Capacity	-0.059*	-0.068*	-0.039*	X	-7.925*	14.94*	-10.32*
Oil shortage signal							
T11: Economic Growth	X	0.029	0.066*	14.97*	X	X	-3.244
Energy policy signal							
T36: Energy Policy	-0.068*	-0.059*	-0.028*	-5.044	-10.76*	37.53*	-19.44*
Others signal							
T7: Physical/Financial Interaction	-0.032	-0.026	-0.018	8.366*	X	X	-1.602
T23: Kuwait Cooperation	-0.0002	X	X	-2.744	9.399*	X	2.056
T30: Energy Outlook	X	X	X	2.111	15.89*	-17.89*	X
T35: World Economy	-0.012	X	X	-8.468*	-2.538	25.98*	-15.39*
Private signals							
Consensus 3 months (SD)	0.311*	0.129*	0.058*	29.62*	-7.273	X	17.29*
Consensus 12 months (SD)	X	X	X	X	-17.99*	42.28*	-54.42*
Control variables							
Production Decision: Cut	0.223*	0.152*	0.152*	-3.488	X	X	X
Production Decision: Increase	0.0005	0.018	X	X	X	X	-3.982
Extraordinary Meetings	X	X	X	X	X	X	X
OPEC Cooperation	X	0.052*	0.052	-48.23*	X	46.88*	-8.494*

Note: This table reports the estimated coefficients of the penalized bootstrap Lasso + partial ridge regressions for both price volatility (in percentage points) and trading positions organized as clusters (see communities in Figure 3). We only report statistically significant topics. $*$ indicates significance at the 5% level. "X" indicates zero value coefficients. To save space, coefficients of trading positions are divided by 1000.

Table 13: Bootstrap de-sparsed Lasso estimation

Signals	1-month	6-month	12-month	P/M/D/P/U	MM	SD	OR
Public signals							
Supply adju. signal							
T5: Rebalancing Market	X	0.029	0.023	-7.403*	X	11.16	X
T9: Cooperation	X	0.054*	0.033	15.83*	X	-32.72*	16.20*
T10: OPEC Producers	X	X	0.0052	5.145	X	-9.743	7.371*
T20: Prod Adjustment/COVID	0.153*	0.043	0.035	3.647	X	-23.46*	16.00*
T25: OPEC Prod. Adjust.	X	X	X	2.377	18.35*	-26.95*	1.346
T29: Minister Energy Negotiations	0.090*	0.017	0.016	-6.875	-5.415	23.42*	-1.999
T34: OPEC/NOPEC Prod. Part.	-0.028	-0.013	X	4.243	21.54*	-43.75*	19.55*
OPEC internal. relations signal							
T15: OPEC-Russia Relations	X	X	X	12.62*	X	X	X
T33: Oil Industry	X	X	X	-7.955	14.77*	X	X
Price vol. & Mk stability signal							
T1: Extraodinary Meetings	X	-0.022	X	-2.485	-9.894*	27.87*	-6.334
T4: Oil Uncertainty/Volatility	-0.043*	-0.043*	-0.069*	X	-15.17*	45.76*	-20.77*
T14: Oil Mkt Stability	0.342*	0.179*	0.119*	-7.131*	-2.202	5.667	X
Prod. Ceiling & LT inv. signal							
T6: LT Strategy	X	X	-0.0025	-1.014	-2.162	16.23*	-8.490*
T21: Adequate Supply	X	X	-0.061*	X	-13.41*	29.08*	-11.30*
T38: Petroleum Industry	X	X	-0.031*	-3.634	-9.307*	29.85*	-8.688*
T40: Global Prod. Capacity	-0.040*	-0.065*	-0.033*	X	-6.961*	18.49*	-10.05*
Oil shortage signal							
T11: Economic Growth	X	0.027	0.062*	14.728*	X	X	-3.013
Energy policy signal							
T36: Energy Policy	-0.051*	-0.056*	-0.021*	-4.681	-9.609*	42.66*	-19.07*
Others signal							
T7: Physical/Financial Interaction	-0.016	-0.023	-0.012	8.083*	X	X	-1.329
T23: Kuwait Cooperation	X	X	X	-2.506	8.605*	X	1.885
T30: Energy Outlook	X	X	X	1.837	14.95*	-19.58*	X
T35: World Economy	X	X	X	-8.188*	-1.468	30.08*	-15.13*
Private signals							
Consensus 3 months (SD)	0.292*	0.126*	0.053*	29.07*	-5.832	X	15.66*
Consensus 12 months (SD)	X	X	X	X	-20.00*	38.43*	-53.31*
Control variables							
Production Decision: Cut	0.214*	0.150*	0.150*	-2.978*	X	X	X
Production Decision: Increase	X	0.014	X	X	X	X	-3.850
Extraordinary Meetings	X	X	X	X	X	X	X
OPEC Cooperation	X	0.049*	0.047*	-48.16*	X	48.98*	-8.398*

Note: This table reports the estimated coefficients of the penalized bootstrap de-parsed Lasso regressions for both price volatility (in percentage points) and trading positions organized as clusters (see communities in Figure 3). We only report statistically significant topics. Estimations and confidence intervals are computed over B = 5000 bootstrap replications. A 5-fold cross validation procedure has been performed to select the Lasso penalty λ. $*$ indicates significance the at 5% level. "X" indicates zero value coefficients. To save space, coefficients of trading positions are divided by 1000.

Table 14: Bayesian Lasso estimation

Signals	1-month	6-month	12-month	P/M/D/P/U	MM	SD	OR
Public signals							
Supply adju. signal							
T5: Rebalancing Market	X	0.0046	0.0081	-3.303*	X	X	X
T9: Cooperation	X	0.046*	0.026	16.09*	X	-34.70*	17.47*
T10: OPEC Producers	X	X	X	X	X	X	5.611*
T20: Prod Adjustment/COVID	0.166*	0.020	0.021	X	X	-22.20*	16.40*
T25: OPEC Prod. Adjust.	X	X	X	X	19.84*	-27.56*	X
T29: Minister Energy Negotiations	0.106*	X	X	-2.334	X	20.96*	X
T34: OPEC/NOPEC Prod. Part.	X	X	X	X	23.46*	-47.75*	20.23*
OPEC internal. relations signal							
T15: OPEC-Russia Relations	X	X	X	13.49*	X	X	X
T33: Oil Industry	X	X	X	-4.258	15.59*	X	X
Price vol. & Mk stability signal							
T1: Extraodinary Meetings	X	X	X	X	-6.492*	25.47*	-0.161
T4: Oil Uncertainty/Volatility	-0.021*	0.031*	-0.081*	X	-14.72*	48.09*	-22.06*
T14: Oil Mkt Stability	0.366*	0.190*	0.136*	-4.181*	X	X	X
Prod. Ceiling & LT inv. signal							
T6: LT Strategy	X	X	X	X	X	9.900*	-7.762*
T21: Adequate Supply	X	X	-0.072*	X	-12.59*	31.46*	-11.12*
T38: Petroleum Industry	X	X	0.026	X	-6.632*	30.73*	-6.236*
T40: Global Prod. Capacity	-0.020*	-0.072*	-0.041*	X	-2.499*	13.43*	-9.717*
Oil shortage signal							
T11: Economic Growth	X	0.056	0.074*	15.57*	X	X	X
Energy policy signal							
T36: Energy Policy	-0.026*	-0.053*	-0.0081*	X	-6.551*	44.68*	-19.54*
Others signal							
T7: Physical/Financial Interaction	X	X	X	6.875*	X	X	X
T23: Kuwait Cooperation	X	X	X	X	7.234*	X	X
T30: Energy Outlook	X	X	X	X	15.62*	-16.43*	X
T35: World Economy	X	X	X	-5.481*	X	31.41*	-15.37*
Private signals							
Consensus 3 months (SD)	0.344*	0.141*	0.064*	29.02*	-3.662	-2.136	25.97*
Consensus 12 months (SD)	X	X	X	X	-21.15*	49.13*	-64.92*
Control variables							
Production Decision: Cut	0.232*	0.165*	0.167*	-2.900*	X	X	X
Production Decision: Increase	X	X	X	X	X	X	X
Extraordinary Meetings	X	X	X	X	X	X	X
OPEC Cooperation	X	0.051*	0.057*	-50.19*	X	55.24*	-6.328*

Note: This table reports the Bayesian lasso estimation for both price volatility (in percentage points) and trading positions organized as clusters (see communities in Figure 3). "X" indicates zero value coefficients. $*$ indicates significance in a bayesian sense. To save space, coefficients of trading positions are divided by 1000.