News Media and Attention Spillover across Energy Markets
A Powerful Predictor of Crude Oil Futures Prices

Cepni, Oguzhan; Nguyen, Duc Khuong ; Sensoy, Ahmet

News Media and Attention Spillover across Energy Markets: A Powerful Predictor of Crude Oil Futures Prices

Oguzhan Cepni[a,b], Duc Khuong Nguyen[c,d,⋆], Ahmet Sensoy[e]

[a]Copenhagen Business School, Copenhagen, Denmark. Email: oce.eco@cbs.dk
[b]Central Bank of the Republic of Turkey, Markets Department, Ankara, Turkey.
[c]IPAG Business School, Paris, France. Email: duc.nguyen@ipag.fr
[d]International School, Vietnam National University, Hanoi, Vietnam
[e]Bilkent University, FBA, Ankara, Turkey. Email: ahmet.sensoy@bilkent.edu.tr
[⋆]Corresponding author: 184 Boulevard Saint Germain, 75006 Paris, France.

Abstract

We develop two news-based investor attention measures from the news trends function of the Bloomberg terminal and investigate their predictive power for returns on crude oil futures contracts with various maturities. Our main results after controlling for relevant macroeconomic variables show that the Oil-based Institutional Attention Index is useful in predicting oil futures returns, especially during price downturn periods, while the forecasting accuracy is further improved when the Commodity Market Institutional Attention Index is used. This forecasting accuracy decreases, however, with the maturity of oil futures contracts. Moreover, we find some evidence of Granger-causality and regime-dependent interactions between investor attention measures and oil futures returns. Finally, variable selection algorithms matter before making predictions since they create the best forecasting results in many cases considered. These findings are important for informed traders and policymakers to better understand the price dynamics of the oil markets.

Keywords: Crude oil returns; Density forecasting; investor attention; time-varying Granger causality; variable selection.

JEL Classification: C51, C53, C58, G17, Q47.

1. Introduction

Crude oil is undoubtedly one of the most important commodities globally as it goes into the production of gasoline, jet fuel as well as many other petroleum products and chemicals. As an essential source of energy, it has also played a prominent role in economic activities of nations worldwide. A positive shock in the oil prices can actually induce an increased cost of production and services, thus raising overall price levels and affecting consumers' purchasing power (Myers et al., 2018; Geiger and Scharler, 2019; Pal and Mitra, 2019). Such an effect might create concerns regarding productivity in the future, lead to negative market sentiment, and decreases in companies' equity values (see, e.g., Cunado and de Gracia (2014); Joo and Park (2021) and references therein). Further impacts can be seen on the GDP since radical changes in oil prices can create a wealth transfer between oil-importing and oil-exporting countries through a shift in terms of trade (see, e.g., Elekdag et al. (2008)). Therefore, the question about oil price driving factors has attracted considerable attention since a reliable forecast of oil prices is of great interest to energy companies, portfolio managers, and policymakers.

To date, several economic and financial variables have been explored for accurate prediction of oil prices, and the related literature can be divided into two categories. The first strand of literature relies extensively on macroeconomic variables, including the U.S. and global macroeconomic aggregates (Lardic and Mignon, 2008; Zagaglia, 2010), economic activity (Kilian and Hicks, 2013), fluctuations in exchange rates (Wu et al., 2012; Brahmasrene et al., 2014; Jawadi et al., 2019; Karlsson et al., 2020), stock market indicators (Narayan and Sharma, 2011; Cunado and de Gracia, 2014; Ye et al., 2020), interest rates (Arora and Tanner, 2013), and policy uncertainty (Li et al., 2016; Garnier and Madlener, 2016; Prokopczuk et al., 2020). It is important to emphasize that oil prices would be mainly affected through the mechanisms related to macroeconomic fundamentals or tangible real factors such as those listed here when commodity markets were not financialized. In this case, the impact of an oil shock hitting the markets would be limited and contained within inflationary consequences.

With the financialization of oil commodities since 2004, the impact of oil prices has reached up a greater extent, especially in the financial markets. The oil markets' new structure has also changed the underlying forces that drive them to become more and more complex, making oil price forecasting more and more difficult. In particular, expectation-based channels such as investor attention have become major factors for the pricing of commodities in the financial markets (see, e.g., Kou et al. (2018); Prange (2021)). Nowadays, the better are the expectations regarding economic outlook, the higher is the interest for oil in financial markets and thus the higher the oil prices. Similarly, the increased investor attention to crude oil can be considered as a leading signal for the upcoming drastic price movements. On top of these, the last three decades have witnessed the rise of innovative financial instruments through derivatives markets, enabling investor attention to have a more vital role in determining the oil prices.

Given that considering information from fundamentals only is insufficient, the second strand of the literature utilizes some attention measures in energy markets to predict oil prices. For instance, Afkhami et al. (2017) use Google search data to measure attention in energy markets and generate proxies that best represent investor attention for energy price volatility prediction. They find that their Google Search Volume based on keywords related to 90 energy-related keywords is a significant predictor of energy price volatility. Qadan and Nama (2018) show that investor sentiment includes informative information for predicting oil prices and volatility. In line with these studies, Basistha et al. (2015), Li et al. (2015), and Campos et al. (2017) also demonstrate that internet search activity or internet concern is helpful to produce better forecasts of oil prices.

It is equally important to note that investor attention measures using information from social media platforms and Google search activity mostly capture retail investor attention. However, retail investors' capability to collect and process information is extremely limited compared to institutional investors (Ben-Rephael et al., 2017). Given the dominant role of institutional investors, it is crucial to examine the impact of their attention on oil prices.

Building on these views, this study utilizes a new overall commodity market index and an oil-specific investor attention index based on the Bloomberg terminal's news count function to capture the more immediate and relevant market antecedents of oil price fluctuations. To the extent that the majority of Bloomberg terminal users are mainly institutional investors who have both the incentives and financial resources to react quickly to significant news about oil price movements, it is likely that the news appears on the Bloomberg terminal is closely followed by commodity traders.

To implement our study, we employ the news trends function of the Bloomberg terminal and collect the news counts which may have signaling information about oil prices. We then construct an Oil-based Institutional Attention Index (OIAI) from the news counts data and utilize it for predicting oil price returns while controlling for several macroeconomic variables such as economic activity and global crude oil production. Similarly, we also construct a Commodity Market Institutional Attention Index (CMIAI) by extracting the common components of news counts of a large set of commodities using the partial least squares (PLS) approach. With the creation of these two indices, we intend to address three critical questions: i) Does the OIAI allow to predict oil prices better than other economic variables? ii) How does the predictive power of the investor attention indexes change in the short- and long-term horizons? iii) Is the OIAI predictive power superior to that of the CMIAI?

Overall, our paper makes several important contributions to the literature on investor attention and oil price predictability. First, we use a new and direct measure of institutional investor attention. Importantly, this measure can capture a broader set of news that may draw institutional investors' attention, allowing us to examine its role in forecasting oil prices. Because the institutional attention index is broadly analogous to the direct measure of retail attention from Google searches, our results complement the investor attention literature that only focuses on retail attention to predict oil prices. Since Bloomberg terminals are essential in disseminating news to institutional investors, our paper also contributes to the broader literature linking the news media to asset prices. Finally, we provide new empirical evidence on the predictive ability of overall energy markets' attention vs. oil-specific attention index.

The rest of this paper is organized as follows. Section 2 describes the data sources and lists the keywords used to collect news counts. Section 3 introduces the methodological framework we use to construct institutional investor attention indices and to examine the predictive ability of attention indices for oil future returns. Section 4 reports and discusses the empirical results.Section 5 provides some concluding remarks.

2. Data

We use the news trends function of the Bloomberg terminal to collect the news counts including the following terms: 'Crude Oil', 'Carbon & Environmental Markets', 'OPEC', 'Diesel', 'Heating Oil', 'Gasoil', 'Gasoline', 'Liquefied Natural Gas', 'Aluminum', 'Copper', 'Iron Ore', 'Ferro-Alloys', 'Nickel', 'Silver', 'Steel', 'Cattle', 'Coffee', 'Corn', 'Soybeans & Soy Products', 'Sugar', 'Wheat', 'Hogs', 'Cocoa' 'Orange Juice', 'Palm Oil', 'Rice', and 'Canola'. Given that Bloomberg terminals are mainly used by institutional investors working in asset management, banking, and institutional financial services (Ben-Rephael et al., 2017), it is expected that the news appearing on Bloomberg terminal is followed by institutional investors who have greater resources and stronger incentives to pay attention to news quickly.

To control the demand and supply conditions, the global oil production data sourced from the Energy Information Agency (EIA) and the ADS index of Aruoba et al. (2009) which is a daily measure of economic activity based on nine macroeconomic indicators (i.e., drivers of economic growth) are used.[1] Furthermore, we obtain data of crude oil futures contracts for a wide range of maturities including 1, 2, 3, 6, 9, 12-months from Bloomberg terminal.

Our monthly dataset covers January 2012 to December 2020, whereby the starting date is

[1]Data for the ADS index is downloaded from: https://www.philadelphiafed.org/surveys-and-data/real-time-data-research/ads

determined by the data availability of the news counts function on the Bloomberg terminal. All series are adjusted for seasonality (where relevant) and appropriately made stationary by either differencing or log-differencing.

3. Methodological Approach

3.1. Constructing institutional investor attention indices

We use the total number of news including the 'Crude Oil' keyword as the oil market investor attention index (OIAI) since it captures the investor attention directly related to oil-specific market.

As for the CMIAI, we construct it by implementing the partial least squares (PLS) method of Wold (1966) instead of the PCA approach. On this topic, Boivin and Ng (2006) indicate that the principal component analysis (PCA) may perform poorly in predicting the target variable since it might be the case that the estimated factors do not include predictive power for the target variable we want to forecast. Similar to PCA, the PLS method yields dimension reduction by estimating a set of latent factors from a large set of variables, but unlike the PCA, it constructs factors that are specifically valuable for forecasting a given target which is oil futures returns in our case.

In particular, we apply the PLS method following the two-step procedure introduced by Friedman et al. (2001). Firstly, each candidate predictor variable x_j $(j = 1, \ldots, p)$ are standardized to have zero mean and unit variance. Secondly, univariate regression coefficients $\gamma_{1j}\widehat{\gamma_{1j}} = \langle x_j, y \rangle$ are stored for each j. Subsequently, the first PLS direction $z_1 = \sum_j \gamma_{1j}\widehat{\gamma_{1j}} x_j$ is computed as the weighted sum of the original set of predictor variables and vector of univariate regression coefficients. Hence, the estimation of the PLS direction incorporates the degree of relation between oil futures returns and common factors. Afterward, the "target" variable y is regressed on z_1, giving a coefficient θ_1, and then all inputs are orthogonalized with respect to z_1. This process is continued until PLS provides a sequence of $l < p$ orthogonal directions.

Given that the PLS utilizes the oil futures returns to create the directions, its solution path is a non-linear function of oil futures returns. Groen and Kapetanios (2016) suggest that while the PCA finds directions that maximize only the variance of the set of predictor variables, the PLS seeks the directions that have high variance and high correlation with the oil futures returns simultaneously. Specifically, the m^{th} PLS direction $\gamma_m \widehat{\gamma_m}$ solves the following optimization problem:

$$\begin{aligned} &\max_{\alpha} \qquad Corr^2(y, X_\alpha)Var(X_\alpha), \\ &\text{subject to} \quad \|\alpha\| = 1, \quad \alpha' S \gamma_l \widehat{\gamma_l} = 0, \quad l = 1, ..., m-1 \end{aligned} \tag{1}$$

where S represents the sample covariance matrix of the x_j.

We use partial least squares (PLS) to summarize the information contained a broad set of news counts including a wide range of commodities, namely: 'Diesel, Heating Oil, Gasoil', 'Gasoline', 'Liquefied Natural Gas', 'Aluminum', 'Copper', 'Iron Ore', 'Ferro-Alloys', 'Nickel', 'Silver', 'Steel', 'Cattle', 'Coffee', 'Corn', 'Soybeans & Soy Products', 'Sugar', 'Wheat', 'Hogs', 'Cocoa' 'Orange Juice', 'Palm Oil', 'Rice', and 'Canola'. We call these estimated PLS-factors as commodity market institutional investor attention index (CMIAI).

3.2. Forecasting experiment

To examine the predictive ability of investor attention indices for oil futures returns, we carry out factor-augmented predictive regressions based on different model specifications. We utilize both a point and a density forecasting exercises to generate predictions from the different model specifications. The evaluation of forecasts is based on the logarithm of the average predictive likelihoods (APL) for comparing densities and, on the mean square forecast errors (MSFEs) for point forecasts. We employ the 50% of the sample to evaluate out-of-sample forecasts, resulting in a period of 108-h observations where APLs and MSFEs are computed. All model specifications are re-estimated at each step using the information available at time t. The forecast horizons evaluated are h = 1, 2, 3, 6, 9, 12-step ahead forecasts. Furthermore, the mean squared forecast error (MSFE)-adjusted test of Clark and West (2007) is implemented to compare forecast performance relatively to the simple autoregressive (AR) model, which is applicable for nested models.

We implement a set of specifications which enable us to isolate any forecast improvements from the incorporation of different type of investor attention indices. In particular, our forecasting exercise considers the following models:

- **Specification 1**: $r_{t+h}^{(n)} = \mu + \mathcal{L}^p r_t + \beta' Controls_t + \varepsilon_{t+h}$
- **Specification 2**: $r_{t+h}^{(n)} = \mu + \mathcal{L}^p r_t + \beta' Controls_t + \vartheta' OIAI_t + \varepsilon_{t+h}$
- **Specification 3**: $r_{t+h}^{(n)} = \mu + \mathcal{L}^p r_t + \beta' Controls_t + \vartheta' OIAI_t + \theta' CMIAI_t + \varepsilon_{t+h}$
- **Specification 4**: $r_{t+h}^{(n)} = \mu + \mathcal{L}^p r_t + \beta' Controls_t + \vartheta' OIAI_t + \theta' CMIAI_t + \delta' ENVMAR_t + \varepsilon_{t+h}$
- **Specification 5**: $r_{t+h}^{(n)} = \mu + \mathcal{L}^p r_t + \beta' Controls_t + \vartheta' OIAI_t + \theta' CMIAI_t + \delta' ENVMAR_t + \psi' OPEC_t \varepsilon_{t+h}$

where $r_{t+h}^{(n)}$ represents oil futures returns for maturities n = 1, 2, 3, 6, 9, 12 months, $Controls_t$ includes the ADS index and oil production. $OIAI_t$ is the oil market investor attention index. $CMIAI_t$ is the commodity market investor attention index estimated using the PLS approach. $ENVMAR_t$ is the green energy attention index based on news counts including the keyword of 'Carbon & Environmental Markets'. $OPEC_t$ refers to news counts including the keyword 'OPEC'.[2] The lag length p of each specification is chosen based on SIC criteria. While Specification 1 allows us to evaluate the importance of macroeconomic conditions in addition to lags of oil futures returns and constant, Specification types from 2 to 5 are an extension that includes a wide variety of investor attention indices.

3.3. Variable selection algorithms

In addition to factor-augmented predictive regressions, we investigate whether there is an improvement in forecast accuracy obtained from appropriate variable selection algorithms.

[2]We set the number of factors to be four for $CMIAI_t$ since the first four factors explain nearly 50% variation in oil future returns for a given maturity. We also utilize the Bai and Ng (2002) criteria, but it chooses too many factors which deteriorate forecast performance.

The reason is that model and parameter uncertainty may result in an adverse impact on the marginal predictive content of factors that are constructed using finite samples of data (Cepni et al., 2019). Hence, we implement variable selection methods to shrinkage coefficients of some predictors to zero before the construction of forecasts. In particular, we analyze alternative variable selection methods, namely Elastic-Net and Least Absolute Shrinkage Operator (LASSO). Accordingly, we choose indicators from the set of variables for each month that includes control variables, *OIAI*, *ENVMAR*, *OPEC*, and all the four factors of the CMIAI index.

3.3.1. Least Absolute Shrinkage Operator (LASSO)

We apply the LASSO method introduced by Tibshirani (1996). Contrary to the ridge estimator, LASSO puts an ℓ_1-norm penalty on the regression coefficients for possible shrinkage. The LASSO estimator is formulated below:

$$\hat{\beta}^{lasso} = \min_{\beta} \quad \|Y - X\beta\|_2 + \lambda \sum_{j=1}^{N} |\beta_j|, \tag{2}$$

where λ is a tuning parameter that adjusts the strength of the ℓ_1-norm penalty. Considering that our objective function in the LASSO is not differentiable, we employ the efficient iterative algorithm (shooting algorithm) introduced by Fu (1998) for numerical optimization.

3.3.2. Elastic Net (ENET)

Tibshirani (1996) shows that the predictive accuracy of the LASSO approach deteriorates compared to the forecast performance of the ridge regression in the presence of highly correlated variables. To overcome this problem, Zou and Hastie (2005) suggest incorporating a hybrid version of the estimators Ridge regression and LASSO, called the elastic net estimator (ENET). The ENET estimator is denoted as follows:

$$\hat{\beta}^{EN} = \min_{\beta} \quad \|Y - X\beta\|_2 + \lambda_1 \sum_{j=1}^{N} |\beta_j| + \lambda_2 \sum_{j=1}^{N} {\beta_j}^2, \tag{3}$$

where λ_1 and λ_2 are tuning parameters controlling the two penalty functions. Similar to the LASSO, the ENET also yields a possible shrinkage of coefficients to zero.

3.4. Time-varying Granger causality test of Rossi and Wang (2019)

We implement the time-varying parameter robust Granger-causality method (TVP-GC) of Rossi and Wang (2019) to check the causal relationship between oil futures returns and investor attention indices over time for maturities $n = 1, 2, 3, 6, 9, 12$ months. The main advantage of the TVP-GC method is that it is more robust than the standard Granger causality test in the presence of instabilities and allows us to distinguish the periods when Granger causality exists or disappears in the data.[3] Given that our sample includes the substantial oil price drop periods leading to a destabilizing effect on oil futures returns, the TVP-GC method enables us to examine the time-varying causal relationships over time.[4] Hence, it is more optimal to use the TVP-GC method to obtain a more appropriate estimation of the relationship than a constant parameter Granger causality method.

In particular, we consider a VAR model with time-varying parameters as follows:

$$y_t = \Phi_{1,t} y_{t-1} + \Phi_{2,t} y_{t-2} + \cdots + \Phi_{p,t} y_{t-p} + \epsilon_t \tag{4}$$

where $y_t = [y_{1,t}, y_{2,t}, \ldots, y_{k,t}]'$ is an k×1 vector, $\Phi_{j,t}, j = 1, \ldots, p$ are functions of time-varying coefficient matrices, and ϵ_t denotes idiosyncratic shocks which are assumed to be heteroscedastic and serially correlated. The endogenous variables vector y_t in the VAR model includes alternatively oil futures returns for maturities $n = 1, 2, 3, 6, 9, 12$ months, *OIAI*, *ENVMAR*, *OPEC*, and the first common factor of the CMIAI index, the ADS index and the oil production.

We test the null hypothesis that the lags of the CMIAI (and alternatively, OIAI) index do not Granger cause the oil futures returns for a given maturity, where θ_t denotes an

[3] Rossi (2005) suggests that traditional VAR-based test statistics do not yield coherent inferences about statistical significance in case of parameter instability.

[4] On 20 April 2020, WTI Crude futures contracts dropped below \$0 for the first time in history.

appropriate subset of vec $(\Phi_{1,t}, \Phi_{2,t}, \ldots, \Phi_{p,t})$:

$$H_0 : \theta_t = 0, \quad \forall t = 1, 2 \ldots T \tag{5}$$

In doing so, we report test statistics such as the mean Wald (MeanW) test, the exponential Wald (ExpW) test, the Nyblom test, and the Quandt likelihood-ratio (SupLR) test following from Rossi and Wang (2019).[5] The lag length of the VAR model is selected based on the Schwarz Information Criterion (SIC).

4. Empirical Results

4.1. Predictability of oil futures returns: The role of investor attention indices

Our empirical analysis begins with the examination of the contemporaneous impact of investor attention on the oil futures returns. Table 1 presents in-sample estimation results for the entire sample, indicating that attention indices together with the control variables explain 58% of the variation in oil futures returns for the one-month maturity. We also observe that while the estimated coefficient associated with the OIAI is negative and significant at a 1% level, the coefficient of the CMIAI is positive and larger in absolute value than that of the OIAI. This result suggests that the increased oil-specific investor attention implies a downside risk for oil futures returns rather than an upside move. One possible explanation for this result is that investors in the oil market are likely to pay more attention to the bad news related to downside risk (Han et al., 2017). Hence, institutional investors seem to be more worried about falling oil prices rather than rising oil prices. Similar shreds of evidence are also observed in literature for other financial assets (Xiao et al., 2018; Chen et al., 2020a; Dzieliński et al., 2018). On the contrary, the significant positive coefficient

[5] The TVP-GC test is implemented using the gcrobustvar command of STATA as provided by Rossi and Wang (2019). Following the extant structural break literature, we choose the standard trimming parameter as 0.10 since the potential break dates are usually trimmed to omit the beginning and end of the sample period.

of CMIAI suggests that investors' attention on the overall commodity market may indeed be an important channel for driving the oil futures returns through attention spillovers across the oil market. Furthermore, the size of the estimated coefficients for CMIAI and OIAI generally decreases with maturity, possibly reflecting that attention contagions present relatively short-lived effects.

Table 2 presents the point forecasting results where model parameters are estimated recursively on a monthly basis in an expanding window. Whereas the entries in the first row of each panel report the actual MSFEs of the benchmark AR model, all other entries are the MSFEs relative to those of the AR model. Hence, the entries below show a better forecast performance for a given specification type compared to the benchmark AR model. A sequence of h-step forecasts for each month, i.e., $h = 1, 2, 3, 6, 9, 12$ are generated. For the purpose of making interpretation and comparison easier, we highlight the smallest MSFEs in bold.

The results presented in Table 2 reveal several interesting findings. First, point forecasts from the models that include oil-specific and overall commodity market investor attention indices generally yield better prediction accuracy than other models that have control variables and different types of attention indices. In particular, the forecast accuracy gains are higher for short-term forecast horizons ($h = 1, 2, 3$), and a nearly 17% reduction in MSFE relative to the benchmark AR model is obtained if we look at the forecast performance of the Specification 3 for two months oil future returns at $h = 1$. The equivalent reductions in MSFE are also available for other maturities. On the other hand, the forecast gains are generally decreasing in the forecast horizon, suggesting that the effect of institutional investor attention indices is short-lived, and its predictive power deteriorates for longer forecast horizons. Hence, our results are in line with the findings of Mbanga et al. (2019) who document that the effect of attention on sentiment is short-lived for medium and large companies' stock returns, resulting in insignificant predictive power for future stock returns at longer forecast horizons.

Second, recall that there are six different maturities and eight forecast horizons, meaning that we have a total of 48 possible comparisons. Of the various specifications, Specification 3, which includes both oil market-specific and commodity market-specific investor attention indices, performs relatively well as it attains the top rank in 27 out of 48 cases. On the contrary, Specification 2, which includes only the oil-specific institutional investor attention index, is not particularly useful for predicting oil futures returns across all maturities. This finding provides evidence of attention spillovers across commodity markets. In other words, a change in the overall commodity market investor attention incorporates predictive power for oil futures returns due to contagion spillovers of investor attention. Moreover, this result is consistent with the view that investors tend to pay more attention to neighboring assets (Chen et al., 2020b).

Third, adding investor attention related to OPEC news produces better forecasts only for the longer forecast horizon ($h = 12$). Fourth, the models that include investor attention related to green energy markets do not lead to forecast improvement in most cases. Fifth, the variable selection algorithms yield improvement in forecast performance. In particular, the LASSO and ENET approaches are the MSFE-best model in 16 out of 48 cases, suggesting that there is still room for accuracy gains from implementing variable selection algorithms before making predictions. Moreover, the plethora of rejections of the Clark and West (2007) test in Table 2 proves that the forecast accuracy gains from utilizing variable selection algorithms are also statistically significant compared to the benchmark AR model.

Table 3 reports the results for density forecasting. The entries in the Table 3 are average predictive likelihoods, and this measure is reported as a spread from the log APL of the benchmark AR model. Thus, positive values imply a better forecast performance than the benchmark model and vice versa. A closer examination of Table 3 indicates that a similar conclusion also holds for density forecasting performance, where Specification 3 is again the top-performing model. This result demonstrates that investor attention indices help improve predictive density forecasting in comparison with the simple AR benchmark, but the largest gains overall are achieved at shorter forecast horizons. While the models including investor attention related to green energy markets (Specifications 4 and 5) do not lead to point forecast improvement in most cases, there are improvements in density forecasting performance in a limited number of cases, especially for longer forecast horizons. As to the variable selection algorithms, they yield better density forecasts in some cases, suggesting that sufficient penalization of coefficients ensures forecasting gains.

4.2. Time-varying robust Granger causality test results

The TVP-GC test results are presented in Table 4. The first column of Table 4 denotes that the standard constant parameter Granger causality test finds no evidence of causality between the CMIAI index and oil futures returns for all maturities. To put it differently, the null of no-Granger causality from the CMIAI index to oil futures returns cannot be rejected with a significance level of 10%. On the other hand, for all maturities and regardless of the test-statistic considered (ExpW, MeanW, Nyblom, SupLR), there exists a consensus among findings that the CMIAI index Granger-causes the oil futures returns when instabilities are taken into account. It is expected that the constant parameter Granger causality test may lead to flawed inferences since it is unlikely that the assumption of homoscedastic idiosyncratic shocks continues to hold in the presence of instabilities during the elevated oil price volatility. However, the TVP-GC method is specifically designed to capture the instabilities in the parameters arising from the shifts in sign and magnitude across time. Hence, this finding suggests that investor attention to the overall commodity market news cannot be overlooked. Changes in the commodity market investor attention lead to changes in oil futures returns because of the asset co-movements. This finding is in line with the behavioral theory outlined in Veldkamp and Wolfers (2007); Peng and Xiong (2006), highlighting the role of investor attention in determining the dynamic co-movement between asset returns.

On the other hand, the TVP-GC method in Table 4 suggests a significant bi-directional causality between oil futures returns and investor attention indices at the 1% significance level. We interpret these results as an evidence of the feedback effect that changes in investor attention cause changes in oil futures returns, while a change in the oil futures returns can attract the attention of institutional investors on the other way around.

In addition to the findings from the analyses reported in Table 4, we present the whole sequence of the Wald statistics across time in Figures 1 and 2 for maturities 1-month and 12-months oil futures returns, which gives more information on when the Granger-causality occurs.[6] In particular, Panels (c) of Figures 1 and 2 show that the Wald test statistic for a causal relationship between the OIAI index and the oil futures return exceeds the threshold levels during the sample period, and the impact is more pronounced around the last quarter of 2019 after the Saudi Aramco drone attack which put high media attention on oil prices. On the other hand, Panel (d) of Figure 2 demonstrates that the Granger-causality relationship between the OIAI index and the 12-month oil futures returns seems to be strikingly strengthened due to the lingering concerns from the onset of the COVID-19 pandemic which led to a quick deterioration in the growth prospect of economies. Finally, Panels (a) of Figures 1 and 2 suggest that investors do not pay more attention to the overall commodity market when operating in the oil market since the causal relation between the CMIAI and oil futures returns disappear after 2019.

4.3. Asymmetric effects of investor attention on oil futures returns

We augment our econometric analysis with a threshold estimation to examine how the effect of investor attention indices may change in different states of the global risk appetite. The CBOE Crude Oil Volatility Index (OVX) – a forward-looking measure of the volatility implied in the pricing of options on crude oil which is compiled by the Chicago Board Options Exchange – is used as a possible proxy for global risk appetite. Commodity traders closely follow fluctuations in the OVX index. This framework allows us to account for a shift in the oil price volatility that frequently compels investors to revise their risk exposures. A simple two-regime threshold model is employed, but rather than imposing a priori an arbitrary

[6] The results for the 2-month, 3-month, 6-month, and 9-month maturities are presented in Figures A1-A4 of the appendix.

classification scheme, we use the methodology suggested by Hansen (1996) to identify the threshold value endogenously.

In particular, we estimate the following model:

$$\begin{aligned} r_t^{(n)} &= \alpha + \beta_1' OIAI_t + \beta_2' CMIAI_t + \beta_3' ENVMAR_t + \beta_4' OPEC_t + \gamma' controls_t + \varepsilon_t \quad \text{if } OVX \leq \phi \\ r_t^{(n)} &= \alpha + \beta_1' OIAI_t + \beta_2' CMIAI_t + \beta_3' ENVMAR_t + \beta_4' OPEC_t + \gamma' controls_t + \varepsilon_t \quad \text{if } OVX > \phi \end{aligned} \tag{6}$$

where ϕ is the threshold variable and errors ε_t are assumed to be white noise processes. We allow heterogeneous error distributions across breaks. Based on the likelihood ratio test statistics, the 41.42 is chosen threshold value of the OVX index. We call these regimes high or low volatility periods where the OVX index is above or below the threshold value.

Table 5 shows the results of threshold regressions. The findings clearly show that the effects of investor attention on oil futures returns are regime-dependent. Although none of the attention indices is statistically significant during times of low volatility (except for 12-month oil futures returns), the attention indices regarding the CMIAI, OIAI and OPEC are statistically significant in most of the cases in high volatility regime. Furthermore, the coefficient of the attention indices decreases in magnitude as the maturity of the oil futures contract increases, indicating that investors pay more attention to short-term price movements. Interestingly, the coefficient of the oil-specific investor attention index is negative and statistically significant. This result confirms the findings reported in Section 4.1 and suggests that investors pay more attention to bad news related to oil prices, which in turn leads to negative changes in the future oil prices through the volatility feedback effect channel (Qadan and Nama, 2018; Xiao and Wang, 2021). That is, negative events are more closely followed by investors than potential gains surrounding positive events in high volatility regime. This result is in line with the findings of previous studies in behavioral finance, providing empirical evidence that investors increase their news searches in response to increased volatility (Abbas et al., 2013; Vozlyublennaia, 2014; Qadan and Nama, 2018).

On the other hand, the coefficient of the overall commodity investor attention index is positive and statistically significant at the 1% level across all maturities. The reason might be that high investor attention towards the overall commodity market reflects the cross-market interdependence and may result in co-movement between asset returns. Similarly, the coefficients of investor attention to OPEC news are statistically significant and positive. This finding suggests that disagreement between OPEC members about current or future production levels may cause oil price increases in the short-term period.

5. Conclusion

To the extent that oil is arguably the most crucial commodity worldwide and remains as a strategic resource, it is essential for investors and policymakers to forecast oil prices with an acceptable error margin. This goal is, however, not simple as it used to be in the 20th century because financial globalization and the ongoing financialization of commodity markets present new challenges to understand the factors that determine oil prices.

In this study, we focused on investor attention as a factor that can be used to predict monthly oil prices. Unlike the earlier studies that consider retail investor attention using metrics from Google search volume or Twitter mentions, we are interested in the attention by institutional investors who are regarded as informed traders. In this respect, using Bloomberg terminal's news count function, we create direct measures of institutional investor attention for (i) crude oil, (ii) carbon & environmental markets, (iii) OPEC, and (iv) the overall commodity markets, and then examine their predictive power over the classical fundamental macro-variables.

Our results show that crude oil investor attention measures we developed are useful in predicting crude oil futures returns, especially during price downturn periods. As to the overall commodity market investor attention, it is found to improve the forecast accuracy obtained from using solely crude oil investor attention. The forecasting accuracy decreases, however, as the maturity of the oil futures increases, showing that these attention measures should be used for short-term forecasts, i.e., a few months. We further consider other attention measures based on terms related to OPEC and green energy markets, but do not find any consistent significant improvement in the forecast accuracy. Another important finding is that variable selection algorithms matter before making predictions since they produce the best forecasting results in most cases considered.

We also execute a dynamic Granger causality analysis on crude oil futures returns and our investor attention indices, and find that the commodity market attention index Granger causes oil futures returns when periodic instabilities are considered. This indicates that investor attention to commodity markets leads to changes in oil future returns because of the systematic price co-movement in commodities. Furthermore, the time-varying causality analysis reveals that oil returns cause our attention indices, suggesting a feedback effect between investor attention and crude oil returns. Therefore, investors and policymakers should follow both variables to better understand the pricing dynamics in the market.

The results from the regime-switching regression analysis show that the effects of investor attention on oil futures returns are state-dependent. In particular, attention indices have an impact on oil returns when the global oil market is in a turbulent period, whereas the significance of this impact weakens or disappears in normal times. This indicates that our investor attention measures could reflect behavioral biases and reactions in the market, and they should be closely monitored, especially in high uncertainty regimes.

This paper contributes to the broader literature linking the news media to asset prices. Our news counts-based investor attention indices can capture a broader set of news that may draw institutional investors' attention, allowing us to examine its role in forecasting oil future returns. Future research can extend our study in several ways. First, an improvement could be implemented by analyzing the news content in Bloomberg to give a score and direction to news. Second, in complementing the forecasting approaches we used in this paper (LASSO regression and Elastic-Net), alternative methodological approaches can be considered for an exhaustive comparison of the obtained results. Finally, one could compare the predictive power of our attention measures by using alternative data & news vendors or even combine them to investigate whether the forecasting accuracy effectively improves or not.

References

Abbas, A.E., Bakır, N.O., Klutke, G.A., Sun, Z., 2013. Effects of risk aversion on the value of information in two-action decision problems. Decision Analysis 10, 257–275.

Afkhami, M., Cormack, L., Ghoddusi, H., 2017. Google search keywords that best predict energy price volatility. Energy Economics 67, 17–27.

Arora, V., Tanner, M., 2013. Do oil prices respond to interes rates? Energy Economics 36, 546–555.

Aruoba, S.B., Diebold, F.X., Scotti, C., 2009. Real-time measurement of business conditions. Journal of Business & Economic Statistics 27, 417–427.

Bai, J., Ng, S., 2002. Determining the number of factors in approximate factor models. Econometrica 70, 191–221.

Basistha, A., Kurov, A., Wolfe, M.H., 2015. Forecasting commodity price volatility with internet search activity. NCCC-Working Paper 134, –.

Ben-Rephael, A., Da, Z., Israelsen, R.D., 2017. It depends on where you search: Institutional investor attention and underreaction to news. Review of Financial Studies 30, 3009–3047.

Boivin, J., Ng, S., 2006. Are more data always better for factor analysis? Journal of Econometrics 132, 169–194.

Brahmasrene, T., Huang, J.C., Sissoko, Y., 2014. Crude oil prices and exchange rates: Causality, variance decomposition and impulse response. Energy Economics 44, 407–412.

Campos, I., Cortazar, G., Reyes, T., 2017. Modeling and predicting oil VIX: Internet search volume versus traditional variables. Energy Economics 66, 194–204.

Cepni, O., Güney, I.E., Swanson, N.R., 2019. Nowcasting and forecasting gdp in emerging markets using global financial and macroeconomic diffusion indexes. International Journal of Forecasting 35, 555–572.

Chen, S., Zhang, W., Feng, X., Xiong, X., 2020a. Asymmetry of retail investors' attention and asymmetric volatility: Evidence from china. Finance Research Letters 36, 101334.

Chen, X., An, L., Yu, J., 2020b. Attention spillover in asset pricing. Available at SSRN .

Clark, T.E., West, K.D., 2007. Approximately normal tests for equal predictive accuracy in nested models. Journal of Econometrics 138, 291–311.

Cunado, J., de Gracia, F.P., 2014. Oil price shocks and stock market returns: Evidence for some European countries. Energy Economics 42, 365–377.

Dzieliński, M., Rieger, M.O., Talpsepp, T., 2018. Asymmetric attention and volatility asymmetry. Journal of Empirical Finance 45, 59–67.

Elekdag, S., Lalonde, R., Laxton, D., Muir, D., Pesenti, P., 2008. Oil price movements and the global economy: A model-based assessment. IMF Staff Papers 55, 297–311.

Friedman, J., Hastie, T., Tibshirani, R., et al., 2001. The elements of statistical learning. volume 1. Springer Series in Statistics New York.

Fu, W.J., 1998. Penalized regressions: the bridge versus the lasso. Journal of Computational and Graphical Statistics 7, 397–416.

Garnier, E., Madlener, R., 2016. The influence of policy regime risks on investments in innovative energy technology. Energy Journal 37, –.

Geiger, M., Scharler, J., 2019. How do consumers assess the macroeconomic effects of oil price fluctuations? evidence from us survey data. Journal of Macroeconomics 62, 103134.

Groen, J.J., Kapetanios, G., 2016. Revisiting useful approaches to data-rich macroeconomic forecasting. Computational Statistics & Data Analysis 100, 221–239.

Han, L., Lv, Q., Yin, L., 2017. Can investor attention predict oil prices? Energy Economics 66, 547–558.

Hansen, B.E., 1996. Inference when a nuisance parameter is not identified under the null hypothesis. Econometrica: Journal of the Econometric Society , 413–430.

Jawadi, F., Louhichi, W., Ameur, H.B., Ftiti, Z., 2019. Do jumps and co-jumps improve volatility forecasting of oil and currency markets? Energy Journal 40, –.

Joo, Y.C., Park, S.Y., 2021. The impact of oil price volatility on stock markets: Evidences from oil-importing countries. Energy Economics , 105413.

Karlsson, H.K., Mansson, K., Sjolander, P., 2020. Unveiling the time-dependent dynamics between oil prices and exchange rates: A wavelet-based panel analysis. Energy Journal 41, –.

Kilian, L., Hicks, B., 2013. Did unexpectedly strong economic growth cause the oil price shock of 2003–2008? Journal of Forecasting 32, 385–394.

Kou, Y., Ye, Q., Zhao, F., Wang, X., 2018. Effects of investor attention on commodity futures markets. Finance Research Letters 25, 190–195.

Lardic, S., Mignon, V., 2008. Oil prices and economic activity: An asymmetric cointegration approach. Energy Economics 30, 847–855.

Li, L., Yin, L., Zhou, Y., 2016. Exogenous shocks and the spillover effects between uncertainty and oil price. Energy Economics 54, 224–234.

Li, X., Ma, J., Wang, S., Zhang, X., 2015. How does Google search affect trader positions and crude oil prices? Economic Modelling 49, 162–171.

Mbanga, C., Darrat, A.F., Park, J.C., 2019. Investor sentiment and aggregate stock returns: the role of investor attention. Review of Quantitative Finance and Accounting 53, 397–428.

Myers, R.J., Johnson, S.R., Helmar, M., Baumes, H., 2018. Long-run and short-run relationships between oil prices, producer prices, and consumer prices: What can we learn from a permanent-transitory decomposition? The Quarterly Review of Economics and Finance 67, 175–190.

Narayan, P.K., Sharma, S.S., 2011. New evidence on oil price and firm returns. Journal of Banking and Finance 35, 3253–3262.

Pal, D., Mitra, S.K., 2019. Asymmetric oil price transmission to the purchasing power of the us dollar: A multiple threshold nardl modelling approach. Resources Policy 64, 101508.

Peng, L., Xiong, W., 2006. Investor attention, overconfidence and category learning. Journal of Financial Economics 80, 563–602.

Prange, P., 2021. Does online investor attention drive the co-movement of stock-, commodity-, and energy markets? insights from google searches. Energy Economics 99, 105282.

Prokopczuk, M., Simen, C.W., Wichmann, R., 2020. The natural gas announcement day puzzle. Energy Journal 42, –.

Qadan, M., Nama, H., 2018. Investor sentiment and the price of oil. Energy Economics 69, 42–58.

Rossi, B., 2005. Optimal tests for nested model selection with underlying parameter instability. Econometric Theory , 962–990.

Rossi, B., Wang, Y., 2019. Vector autoregressive-based granger causality test in the presence of instabilities. The Stata Journal 19, 883–899.

Tibshirani, R., 1996. Regression shrinkage and selection via the lasso. Journal of the Royal Statistical Society: Series B (Methodological) 58, 267–288.

Veldkamp, L., Wolfers, J., 2007. Aggregate shocks or aggregate information? costly information and business cycle comovement. Journal of Monetary Economics 54, 37–55.

Vozlyublennaia, N., 2014. Investor attention, index performance, and return predictability. Journal of Banking & Finance 41, 17–35.

Wold, H., 1966. Estimation of principal components and related models by iterative least squares. Multivariate Analysis , 391–420.

Wu, C.C., Chung, H., Chang, Y.H., 2012. The economic value of co-movement between oil price and exchange rate using copula-based GARCH models. Energy Economics 34, 270–282.

Xiao, J., Wang, Y., 2021. Investor attention and oil market volatility: Does economic policy uncertainty matter? Energy Economics 97, 105180.

Xiao, J., Zhou, M., Wen, F., Wen, F., 2018. Asymmetric impacts of oil price uncertainty on chinese stock returns under different market conditions: Evidence from oil volatility index. Energy Economics 74, 777–786.

Ye, Z., Hu, C., He, L., Ouyang, G., Wen, F., 2020. The dynamic time-frequency relationship between international oil prices and investor sentiment in China: A wavelet coherence analysis. Energy Journal 41, –.

Zagaglia, P., 2010. Macroeconomic factors and oil futures prices: A data-rich model. Energy Economics 32, 409–417.

Zou, H., Hastie, T., 2005. Regularization and variable selection via the elastic net. Journal of the Royal Statistical Society: series B (statistical methodology) 67, 301–320.

Table 1: In-sample regressions of monthly oil future returns on attention indices across maturities

	$r_t^{(1)}$	$r_t^{(2)}$	$r_t^{(3)}$	$r_t^{(6)}$	$r_t^{(9)}$	$r_t^{(12)}$
$CMIAI$	0.372***	0.356***	0.323***	0.286***	0.323***	0.286***
	(0.090)	(0.084)	(0.083)	(0.072)	(0.066)	(0.061)
$OIAI$	-0.306***	-0.249***	-0.209***	-0.183***	-0.209***	-0.183
	(0.090)	(0.066)	(0.060)	(0.054)	(0.049)	(0.045)
$OPEC$	0.051**	0.051**	0.045**	0.039**	0.045**	0.039**
	(0.026)	(0.024)	(0.023)	(0.020)	(0.017)	(0.016)
$ENVMAR$	0.014	0.008	0.003	-0.002	0.003	-0.002
	(0.025)	(0.023)	(0.022)	(0.019)	(0.018)	(0.017)
ADS	0.019***	0.011***	0.008***	0.006***	0.008***	0.006***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)
$OILPROD$	-0.255	-0.167	-0.108	-0.103	-0.108	-0.103
	(0.184)	(0.124)	(0.112)	(0.100)	(0.092)	(0.086)
$Constant$	0.001	-0.002	-0.004	-0.004	-0.004	-0.004
	(0.008)	(0.007)	(0.007)	(0.006)	(0.005)	(0.005)
F-statistic	22.64	15.86	12.33	10.82	10.51	10.47
Prob(F-statistic)	0.00	0.00	0.00	0.00	0.00	0.00
R^2	0.576	0.488	0.425	0.394	0.387	0.386

The table reports the estimates from OLS regressions of oil future returns on the variables in columns for maturities n=1, 2, 3, 6, 9, 12-months. Robust standard errors are reported in parentheses. Entries super-scripted with an asterisk denote the statistical significance (*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$).

Table 2: Point forecast performance: MSFEs relative to benchmark AR model

1-month	h=1	h=2	h=3	h=4	h=5	h=6	h=9	h=12
AR	0.086	0.083	0.083	0.086	0.083	0.083	0.084	0.091
Specification -1	1.022	1.071	1.023	1.042	1.029	1.053	1.030	1.016
Specification -2	0.927*	1.147	1.050	1.015	1.082	1.087	1.090	1.036
Specification -3	0.843***	0.916***	0.886***	0.919***	1.024	0.935**	**0.895****	**0.896****
Specification -4	0.851***	0.934***	0.887***	0.923***	1.043	0.941**	0.908**	0.906**
Specification -5	**0.841*****	0.938***	**0.884*****	0.939***	1.059	0.943**	0.940**	0.904**
ENET	0.864***	**0.902*****	0.978***	0.903***	0.991**	0.957**	0.933***	0.920**
LASSO	0.859***	0.917***	0.994	**0.895*****	**0.962*****	**0.929****	0.915***	0.933**
2-months								
AR	0.095	0.092	0.091	0.094	0.093	0.093	0.094	0.102
Specification -1	0.997	1.035	0.996	1.035	1.024	1.052	1.031	1.016
Specification -2	0.930**	1.128	1.009	1.018	1.060	1.068	1.044	1.053
Specification -3	**0.830*****	0.910 ***	**0.837*****	0.930***	0.978***	0.973*	**0.883*****	0.854***
Specification -4	0.842***	0.926***	0.846***	0.934***	0.994*	0.978**	0.899**	0.869***
Specification -5	0.832***	0.929***	0.853***	0.941***	1.005	0.986*	0.923**	**0.838*****
ENET	0.860***	0.902***	0.911***	**0.911*****	**0.970*****	**0.951*****	0.921***	0.922***
LASSO	0.850	**0.894*****	0.901***	0.920***	0.973***	0.955***	0.955***	0.877**
3-months								
AR	0.095	0.096	0.096	0.083	0.081	0.081	0.082	0.089
Specification -1	1.022	1.019	0.992*	1.036	1.012	1.057	1.029	1.015
Specification -2	1.002	1.072	0.991*	1.016	1.064	1.080	1.049	1.057
Specification -3	**0.887*****	**0.906*****	**0.881*****	0.916***	1.002	0.970*	**0.875*****	0.877***
Specification -4	0.898***	0.912***	0.887***	0.923***	1.023	0.979*	0.890**	0.889***
Specification -5	0.899***	0.923***	0.894***	0.934***	1.036	0.989	0.914**	**0.861*****
ENET	0.931***	0.927***	0.963***	0.893***	**0.980*****	0.954***	0.917***	0.922***
LASSO	0.956***	0.935***	1.009	**0.887*****	0.983**	**0.940***	0.948***	0.914***
6-months								
AR	0.084	0.084	0.084	0.089	0.089	0.090	0.091	0.097
Specification -1	1.023	1.022	0.993*	1.023	1.018	1.038	1.019	1.017
Specification -2	1.000	1.067	0.993*	1.008	1.027	1.051	1.036	1.047
Specification -3	**0.889*****	**0.895*****	**0.878*****	0.922***	**0.963****	0.944**	**0.847*****	0.909***
Specification -4	0.896***	0.903***	0.880***	0.932***	0.975**	0.955*	0.857***	0.918**
Specification -5	0.901***	0.913***	0.889***	0.938***	0.980**	0.959*	0.877***	**0.906*****
ENET	0.930***	0.927***	0.965***	**0.904*****	0.972***	0.936***	0.881***	0.915***
LASSO	0.959***	0.955***	0.996**	0.937***	0.968***	**0.932****	0.892**	0.928***
9-months								
AR	0.077	0.076	0.077	0.081	0.081	0.081	0.082	0.088
Specification -1	1.018	1.024	0.986*	1.020	1.014	1.045	1.016	1.012
Specification -2	0.995*	1.069	0.983**	1.006	1.020	1.055	1.032	1.043
Specification -3	**0.879*****	**0.885*****	**0.865*****	**0.910*****	**0.954****	0.942**	**0.838*****	**0.904*****
Specification -4	0.886***	0.891***	0.872***	0.922***	0.970**	0.951**	0.845***	0.911**
Specification -5	0.889***	0.898***	0.882***	0.932***	0.978**	0.956*	0.866***	0.909***
ENET	0.935***	0.910***	0.945***	0.929***	0.971***	**0.922*****	0.895***	0.934***
LASSO	0.910***	0.902***	0.938***	0.922***	0.973***	0.925**	0.899**	0.922***
12-months								
AR	0.071	0.071	0.071	0.074	0.074	0.075	0.075	0.080
Specification -1	1.023	1.023	0.985*	1.023	1.014	1.044	1.022	1.015
Specification -2	0.993*	1.065	0.984**	1.002	1.020	1.053	1.028	1.049
Specification -3	**0.873*****	**0.874*****	**0.863*****	**0.906*****	**0.957****	0.938**	**0.838*****	0.909**
Specification -4	0.880***	0.878***	0.870***	0.918***	0.967**	0.949**	0.850***	0.924**
Specification -5	0.882***	0.886***	0.882***	0.926***	0.980**	0.955*	0.870***	0.921***
ENET	0.912***	0.906***	0.945***	0.945***	0.963***	0.932***	0.921***	**0.902*****
LASSO	0.932***	0.917***	0.964***	0.912***	0.975***	**0.916****	0.927**	0.937***

The entries are MSFEs, with the Specification types that yields the smallest MSFE are highlighted in bold. The entries in the first row correspond to actual point MSFEs of AR model, while all other entries are relative MSFEs relative to the AR model. Hence, a value smaller than one implies that the corresponding specification type produces more accurate forecasts than those of the AR model. Entries marked with an asterisk(s) (*** 1% level; ** 5% level; * 10% level) are significantly superior to the AR model, based on the Clark and West (2007) equal predictive ability test.

Table 3: Density forecast performance: log APLs relative to benchmark AR model

1-month	h=1	h=2	h=3	h=4	h=5	h=6	h=9	h=12
Specification -1	-0.044	-0.036	-0.024	-0.045	-0.026	-0.048	-0.032	-0.049
Specification -2	0.083	-0.094	-0.038	-0.023	-0.073	-0.070	-0.337	-0.062
Specification -3	0.182	**0.149**	**0.114**	0.074	0.000	**0.085**	**0.103**	**0.083**
Specification -4	0.180	0.117	0.114	0.081	-0.040	0.079	0.069	0.077
Specification -5	**0.191**	0.112	0.112	0.073	-0.065	0.068	0.073	0.081
ENET	0.161	0.135	0.005	**0.132**	0.056	0.065	-0.023	0.078
LASSO	0.177	0.146	-0.030	0.126	**0.078**	0.062	0.055	0.065
2-months								
Specification -1	0.004	0.030	0.017	-0.045	0.021	-0.058	-0.043	-0.073
Specification -2	0.089	-0.210	-0.059	-0.002	-0.021	-0.086	-0.031	-0.106
Specification -3	**0.229**	0.107	**0.198**	0.083	0.094	0.024	**0.119**	0.173
Specification -4	0.215	0.156	0.180	0.060	0.052	0.028	0.110	0.166
Specification -5	0.227	0.138	0.173	0.044	0.036	0.000	0.093	**0.214**
ENET	0.194	0.162	0.103	**0.094**	**0.096**	**0.056**	0.026	0.090
LASSO	0.191	**0.167**	0.107	0.063	0.088	0.035	-0.001	0.146
3-months								
Specification -1	**0.193**	-0.021	0.012	-0.036	-0.007	-0.060	-0.037	-0.020
Specification -2	0.030	-0.108	-0.003	-0.027	-0.084	-0.068	-0.275	-0.068
Specification -3	0.157	**0.092**	**0.434**	0.089	-0.008	0.035	0.060	0.139
Specification -4	0.129	0.067	0.238	0.071	-0.060	0.017	**0.113**	0.135
Specification -5	0.134	0.060	0.227	0.054	-0.061	0.002	0.109	**0.162**
ENET	0.129	0.073	0.049	0.113	0.005	0.048	0.077	0.099
LASSO	0.106	0.066	0.010	**0.118**	**0.022**	**0.058**	0.043	0.118
6-months								
Specification -1	0.011	0.105	-0.074	-0.034	0.129	-0.035	0.160	-0.019
Specification -2	0.008	0.096	-0.082	-0.029	0.008	0.531	-0.080	-0.488
Specification -3	0.526	**0.625**	**0.457**	**0.296**	0.280	0.046	0.331	-0.305
Specification -4	0.115	0.131	0.415	0.072	**0.336**	0.426	0.220	-0.174
Specification -5	0.560	0.090	0.339	0.203	0.294	**0.672**	**0.637**	-0.291
ENET	0.518	0.080	0.166	0.149	0.055	0.090	0.462	0.160
LASSO	**0.642**	0.078	-0.063	0.096	-0.042	0.122	0.237	**0.200**
9-months								
Specification -1	0.332	0.040	0.368	-0.133	-0.007	0.035	-0.095	0.497
Specification -2	0.041	0.262	0.270	**0.483**	0.275	-0.475	-0.333	-0.105
Specification -3	**0.712**	0.357	0.717	0.257	**0.464**	0.123	0.035	**0.708**
Specification -4	0.640	0.200	**0.732**	0.072	0.449	0.178	**0.512**	0.445
Specification -5	0.451	-0.055	0.598	0.402	0.449	**0.271**	0.113	0.704
ENET	0.109	**0.467**	0.667	0.014	0.253	-0.314	0.007	0.020
LASSO	0.103	0.382	0.509	0.025	0.099	-0.293	0.089	0.315
12-months								
Specification -1	0.027	-0.292	0.185	-0.068	-0.356	-0.100	-0.196	-0.498
Specification -2	0.276	-0.004	0.120	-0.202	-0.270	-0.212	-0.267	-0.801
Specification -3	0.122	0.176	0.224	0.032	**0.148**	0.139	0.252	-0.302
Specification -4	0.380	**0.349**	0.436	**0.045**	0.071	**0.176**	**0.260**	-0.197
Specification -5	**0.394**	0.139	**0.503**	-0.259	-0.072	0.101	0.250	**0.058**
ENET	-0.085	-0.161	0.457	-0.013	-0.383	0.095	-0.266	-0.607
LASSO	-0.090	-0.138	-0.089	-0.235	0.007	-0.040	-0.059	-0.249

Entries in columns 2-9 of this Table are average predictive likelihoods in logarithms (log APLs) relative to the values of the AR benchmark. Hence, positive log APLs values indicates an improvement relative to the benchmark and vice-versa for negative values. Entries in boldface indicate the best performing model in terms of density forecasting for a given maturity and forecast horizon.

Table 4: Time varying Granger causality test results

1-month	χ^2_q	ExpW	MeanW	Nyblom	SupLR	6-months	χ^2_q	ExpW	MeanW	Nyblom	SupLR
CMIAI $\rightarrow$ Oil	0.12	86.83	53.73	7.54	181.66	CMIAI $\rightarrow$ Oil	0.24	57.06	31.77	14.79	121.65
	(0.73)	(0.00)	(0.00)	(0.00)	(0.00)		(0.62)	(0.00)	(0.00)	(0.00)	(0.00)
Oil $\rightarrow$ CMIAI	0.45	23.86	18.74	116.95	56.59	Oil $\rightarrow$ CMIAI	1.77	219.67	89.52	65.66	448.22
	(0.51)	(0.00)	(0.00)	(0.00)	(0.00)		(0.18)	(0.00)	(0.00)	(0.00)	(0.00)
OIAI $\rightarrow$ Oil	4.29	227.76	169.58	65.73	464.41	OIAI $\rightarrow$ Oil	3.01	233.68	128.79	142.19	476.25
	(0.04)	(0.00)	(0.00)	(0.00)	(0.00)		(0.08)	(0.00)	(0.00)	(0.00)	(0.00)
Oil $\rightarrow$ OIAI	0.04	57.28	18.29	4.45	123.44	Oil $\rightarrow$ OIAI	3.53	509.71	138.78	25.32	1028.31
	(0.85)	(0.00)	(0.00)	(0.00)	(0.00)		(0.06)	(0.00)	(0.00)	(0.00)	(0.00)
2-months						9-months					
CMIAI $\rightarrow$ Oil	0.17	161.16	71.34	27.81	330.68	CMIAI $\rightarrow$ Oil	0.33	41.40	30.98	12.01	91.09
	(0.68)	(0.00)	(0.00)	(0.00)	(0.00)		(0.56)	(0.00)	(0.00)	(0.00)	(0.00)
Oil $\rightarrow$ CMIAI	0.93	191.00	67.83	158.34	390.89	Oil $\rightarrow$ CMIAI	1.31	197.68	72.87	47.33	404.21
	(0.33)	(0.00)	(0.00)	(0.00)	(0.00)		(0.25)	(0.00)	(0.00)	(0.00)	(0.00)
OIAI $\rightarrow$ Oil	2.67	349.31	160.26	85.31	707.50	OIAI $\rightarrow$ Oil	2.82	197.59	121.64	114.43	404.07
	(0.10)	(0.00)	(0.00)	(0.00)	(0.00)		(0.09)	(0.00)	(0.00)	(0.00)	(0.00)
Oil $\rightarrow$ OIAI	2.47	191.74	79.64	30.27	392.23	Oil $\rightarrow$ OIAI	3.18	541.68	127.61	21.85	1092.24
	(0.11)	(0.00)	(0.00)	(0.00)	(0.00)		(0.07)	(0.00)	(0.00)	(0.00)	(0.00)
3-months						12-months					
CMIAI $\rightarrow$ Oil	0.11	98.87	36.77	18.42	206.35	CMIAI $\rightarrow$ Oil	0.40	32.07	28.31	10.41	71.46
	(0.75)	(0.00)	(0.00)	(0.00)	(0.00)		(0.53)	(0.00)	(0.00)	(0.00)	(0.00)
Oil $\rightarrow$ CMIAI	2.76	300.70	137.07	107.51	610.29	Oil $\rightarrow$ CMIAI	1.15	214.98	72.33	35.13	438.82
	(0.09)	(0.00)	(0.00)	(0.00)	(0.00)		(0.28)	(0.00)	(0.00)	(0.00)	(0.00)
OIAI $\rightarrow$ Oil	3.27	345.30	153.42	141.99	699.47	OIAI $\rightarrow$ Oil	2.85	191.35	121.06	98.04	391.57
	(0.07)	(0.00)	(0.00)	(0.00)	(0.00)		(0.09)	(0.00)	(0.00)	(0.00)	(0.00)
Oil $\rightarrow$ OIAI	3.23	328.21	113.82	27.04	665.29	Oil $\rightarrow$ OIAI	2.73	453.56	109.14	18.77	916.00
	(0.07)	(0.00)	(0.00)	(0.00)	(0.00)		(0.09)	(0.00)	(0.00)	(0.00)	(0.00)

Entries correspond to the four alternative test statistics: exponential Wald, mean Wald, Nyblom, and Quandt Likelihood Ratio tests. Similarly, χ^2_q shows the chi-square statistic of constant parameter Granger causality test where the lag length q is selected based on SIC. The corresponding p-values are given in parenthesis. The null hypothesis is that investor attention indices do not Granger cause n-month oil future return where n= 1, 2, 3, 6, 9, 12.

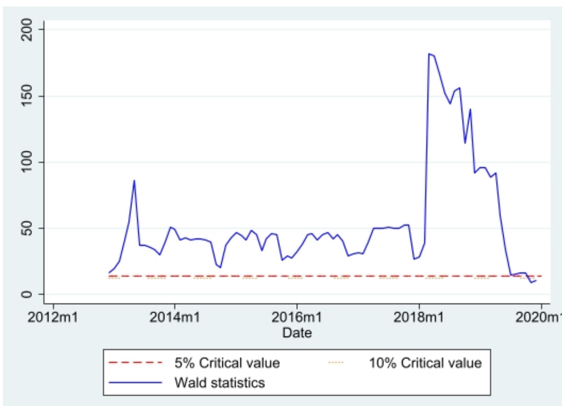
Table 5: Estimation results for high and low volatility regimes across maturities

	$r_t^{(1)}$	$r_t^{(2)}$	$r_t^{(3)}$	$r_t^{(6)}$	$r_t^{(9)}$	$r_t^{(12)}$
			Low Regime (OVX $<$ 41.42)			
$CMIAI$	0.064	0.061	0.049	0.070	0.079	0.082
	(0.072)	(0.073)	(0.074)	(0.067)	(0.061)	(0.056)
$OIAI$	-0.066	-0.065	-0.059	-0.065	-0.070	-0.072*
	(0.067)	(0.065)	(0.061)	(0.053)	(0.047)	(0.043)
$OPEC$	0.008	0.008	0.006	0.008	0.009	0.010
	(0.022)	(0.022)	(0.022)	(0.019)	(0.017)	(0.016)
$ENVMAR$	0.010	0.009	0.010	0.004	0.001	-0.001
	(0.028)	(0.026)	(0.026)	(0.024)	(0.022)	(0.020)
ADS	0.011	0.011	0.011	0.012	0.011	0.010*
	(0.013)	(0.013)	(0.012)	(0.009)	(0.008)	(0.006)
$OILPROD$	-0.032	-0.038	-0.035	-0.060	-0.075	-0.078
	(0.113)	(0.110)	(0.105)	(0.096)	(0.089)	(0.082)
$Constant$	0.008	0.007	0.007	0.006	0.006	0.005
	(0.007)	(0.007)	(0.007)	(0.006)	(0.006)	(0.005)
			High Regime (41.42 $\leq$ OVX)			
$CMIAI$	0.744***	0.723***	0.698***	0.580***	0.509***	0.453***
	(0.147)	(0.144)	(0.166)	(0.149)	(0.138)	(0.132)
$OIAI$	-0.656***	-0.506***	-0.386***	-0.321***	-0.278**	-0.247**
	(0.192)	(0.141)	(0.134)	(0.122)	(0.112)	(0.104)
$OPEC$	0.155*	0.146**	0.119*	0.099*	0.085	0.074
	(0.082)	(0.072)	(0.065)	(0.058)	(0.053)	(0.049)
$ENVMAR$	0.022	0.043	0.045	0.037	0.034	0.031
	(0.060)	(0.048)	(0.048)	(0.043)	(0.041)	(0.040)
ADS	0.012***	0.004*	0.001	0.000	0.000	0.000
	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
$OILPROD$	-0.254	-0.055	0.030	0.011	-0.001	0.016
	(0.399)	(0.295)	(0.296)	(0.288)	(0.273)	(0.258)
$Constant$	-0.030**	-0.040***	-0.042***	-0.041***	-0.040***	-0.039***
	(0.015)	(0.011)	(0.011)	(0.011)	(0.010)	(0.010)
F-statistic	18.03	13.25	10.45	8.70	8.23	8.07
Prob(F-statistic)	0.00	0.00	0.00	0.00	0.00	0.00
R^2	0.72	0.65	0.59	0.55	0.54	0.53

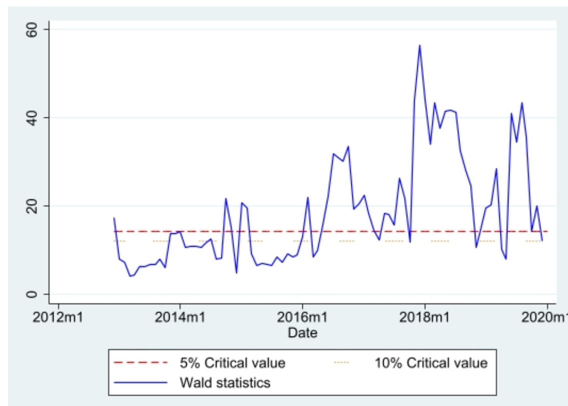
Robust standard errors are reported in parentheses. Entries super-scripted with an asterisk denote the statistical significance (*** p $<$ 0.01, ** p $<$ 0.05, * p $<$ 0.1).

Figure 1: Time varying causality between 1-month oil future return and investor attention indices

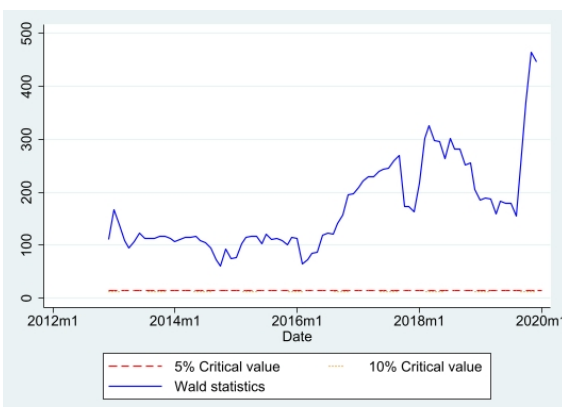
(a) CMIAI → Oil



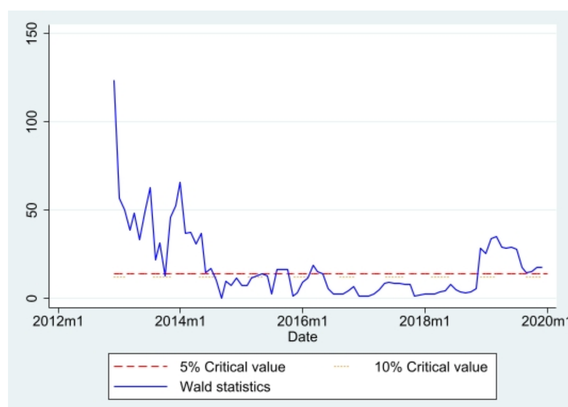
(b) Oil → CMIAI



(c) OIAI → Oil



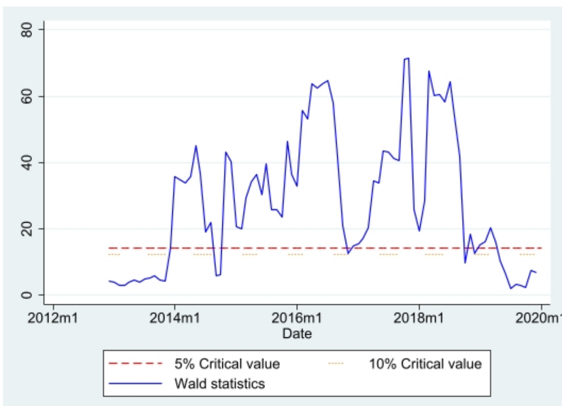
(d) Oil → OIAI



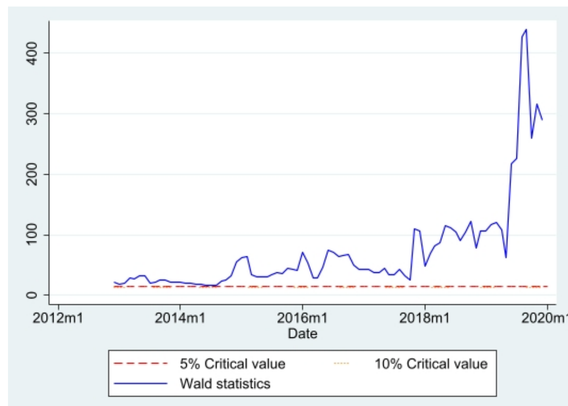
The figure shows Wald statistics of the Granger-causality robust test over time. The null hypothesis is that CMIAI (or OIAI) attention index does not Granger cause return of the 1-month oil future.

Figure 2: Time varying causality between 12-months oil future return and investor attention indices

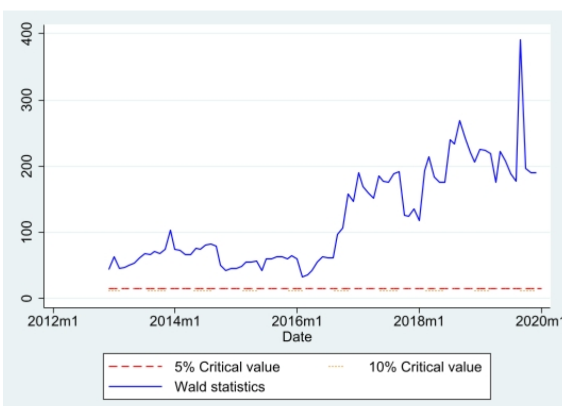
(a) CMIAI → Oil



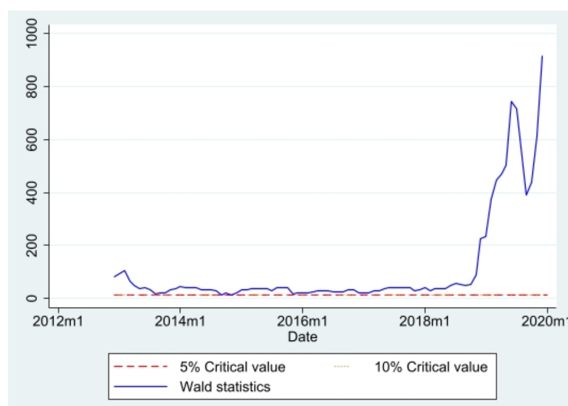
(b) Oil → CMIAI



(c) OIAI → Oil



(d) Oil → OIAI

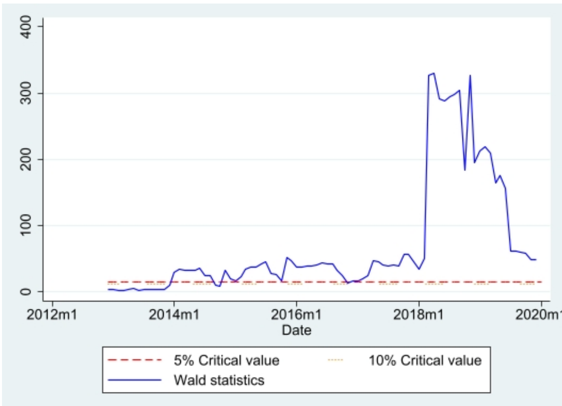


The figure shows Wald statistics of the Granger-causality robust test over time. The null hypothesis is that CMIAI (or OIAI) attention index does not Granger cause return of the 12-months oil future.

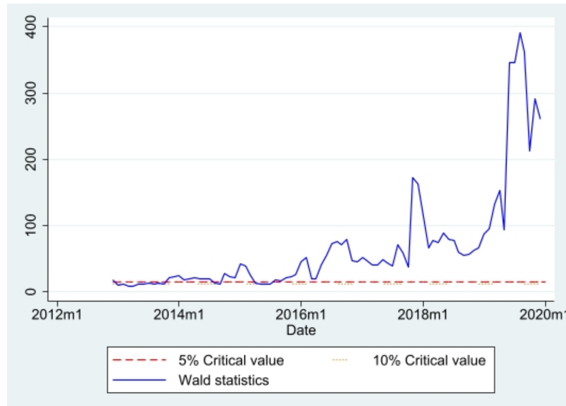
Appendix

Figure A1: Time varying causality between 2-months oil future return and investor attention indices

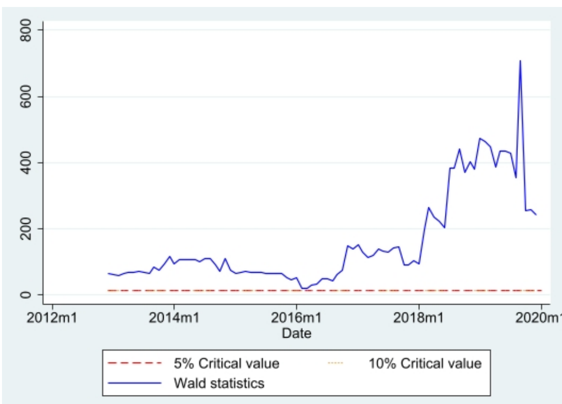
(a) CMIAI → Oil



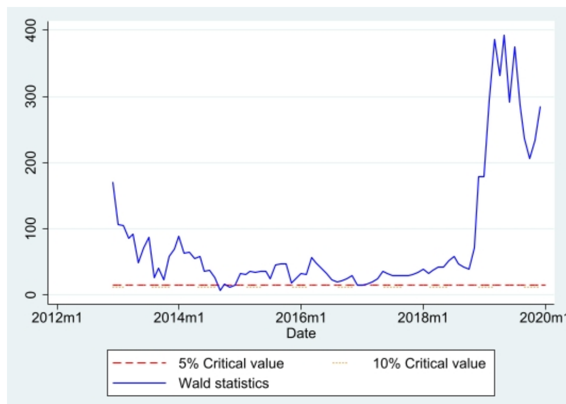
(b) Oil → CMIAI



(c) OIAI → Oil



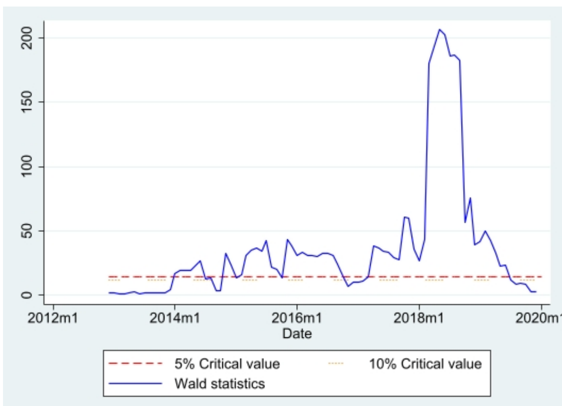
(d) Oil → OIAI



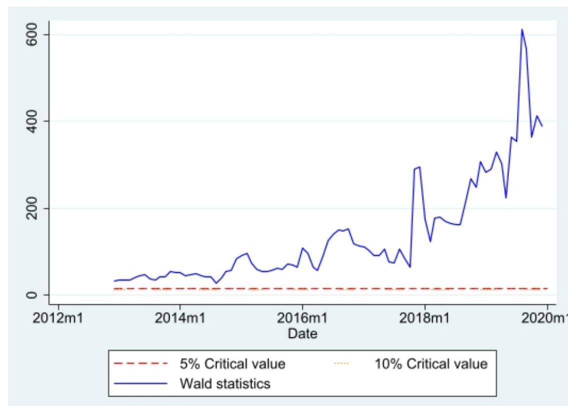
The figure shows Wald statistics of the Granger-causality robust test over time. The null hypothesis is that CMIAI (or OIAI) attention index does not Granger cause return of the 2-months oil future.

Figure A2: Time varying causality between 3-months oil future return and investor attention indices

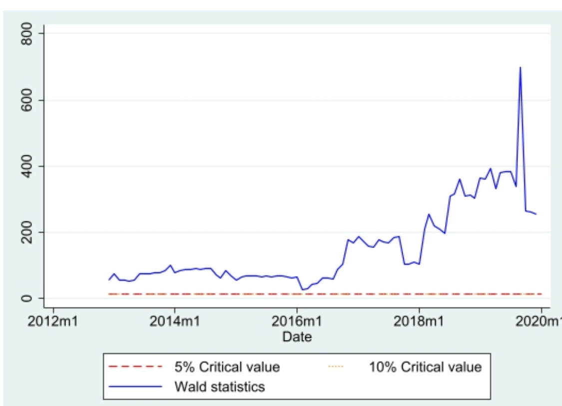
(a) CMIAI → Oil



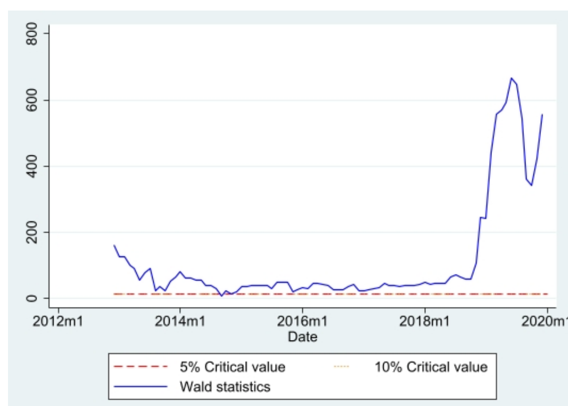
(b) Oil → CMIAI



(c) OIAI → Oil



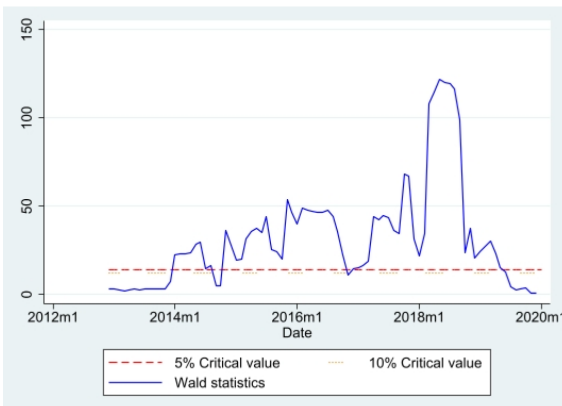
(d) Oil → OIAI



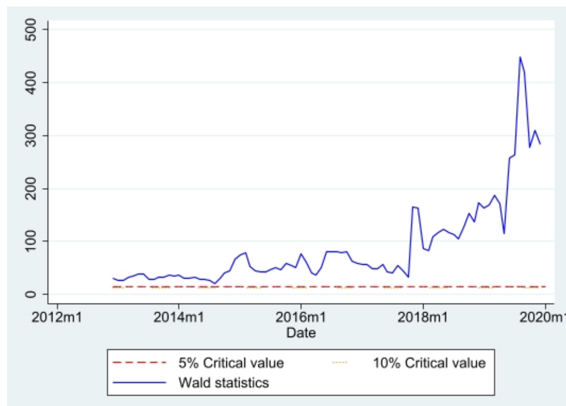
The figure shows Wald statistics of the Granger-causality robust test over time. The null hypothesis is that CMIAI (or OIAI) attention index does not Granger cause return of the 3-months oil future.

Figure A3: Time varying causality between 6-months oil future return and investor attention indices

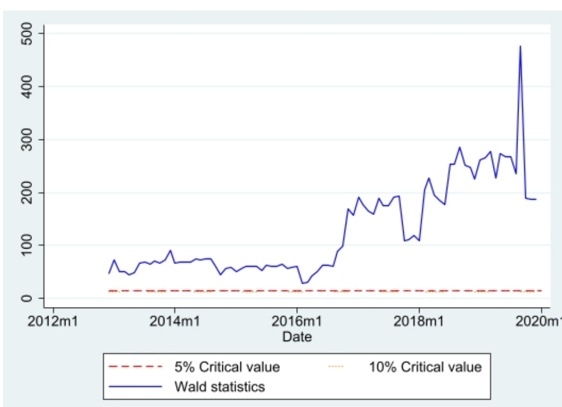
(a) CMIAI → Oil



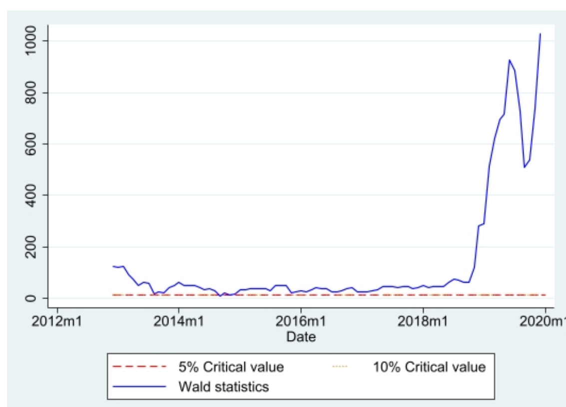
(b) Oil → CMIAI



(c) OIAI → Oil



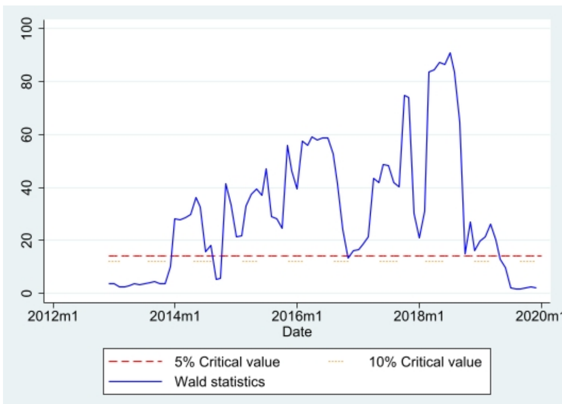
(d) Oil → OIAI



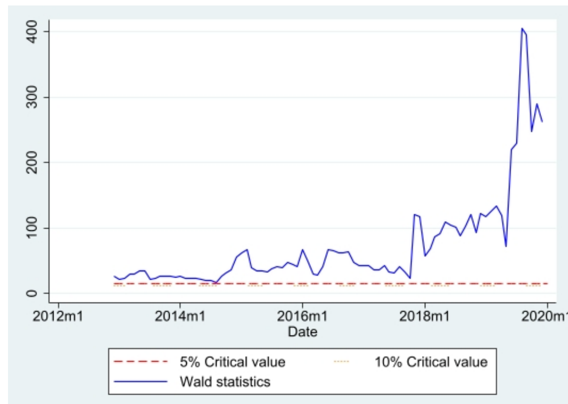
The figure shows Wald statistics of the Granger-causality robust test over time. The null hypothesis is that CMIAI (or OIAI) attention index does not Granger cause return of the 6-months oil future.

Figure A4: Time varying causality between 9-months oil future return and investor attention indices

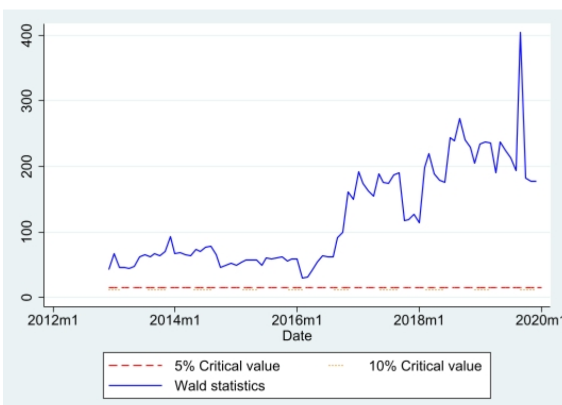
(a) CMIAI → Oil



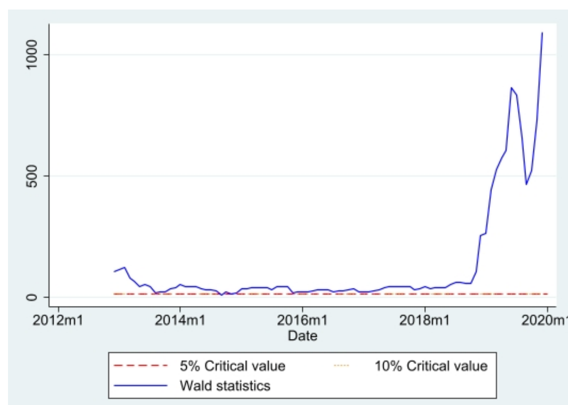
(b) Oil → CMIAI



(c) OIAI → Oil



(d) Oil → OIAI



The figure shows Wald statistics of the Granger-causality robust test over time. The null hypothesis is that CMIAI (or OIAI) attention index does not Granger cause return of the 9-months oil future.