Network

Rimini Centre for Economic Analysis
Working Paper Series

wp 22-13

Oil shocks and investor attention

Georgios Bampinas
Theodore Panagiotidis
Georgios Papapanagiotou

Oil shocks and investor attention

Georgios Bampinas
Panteion University of Athens

Theodore Panagiotidis
University of Macedonia

Georgios Papapanagiotou
University of Macedonia

Abstract

In this paper, we examine the existence of sentiment exposure in oil price returns. We augment the SVAR model of Kilian and Park (*International Economic Review*, 2009, 50, 1267–1287) by including the effects of (1) investors sentiment proxied by Google's search volume index, (2) economic policy uncertainty (EPU) and (3) time variation in both coefficients and the variance-covariance matrix. Our empirical results show that changes in investor attention do exhibit a significant long-lasting impact on oil and stock market returns. Aggregate oil demand and supply shocks have a transitory effect on investor sentiment. We reveal that the impact of EPU is temporary and significant, while EPU responds strongly to shocks on oil prices and stock market returns. In all cases, the magnitude and sign of responses are affected by the timing of the shock. Our findings are robust to an alternative sentiment indicator and once the role of oil inventories is considered.

Keywords: Search Volume Index; investor attention; oil price; stock market; policy uncertainty; time-varying parameter VAR; stochastic volatility; dynamic factor model
JEL Codes: Q02; Q43; Q47; E44; G1; C11

Corresponding author:Theodore Panagiotidis. Email: tpanag@uom.edu.gr

1 Introduction

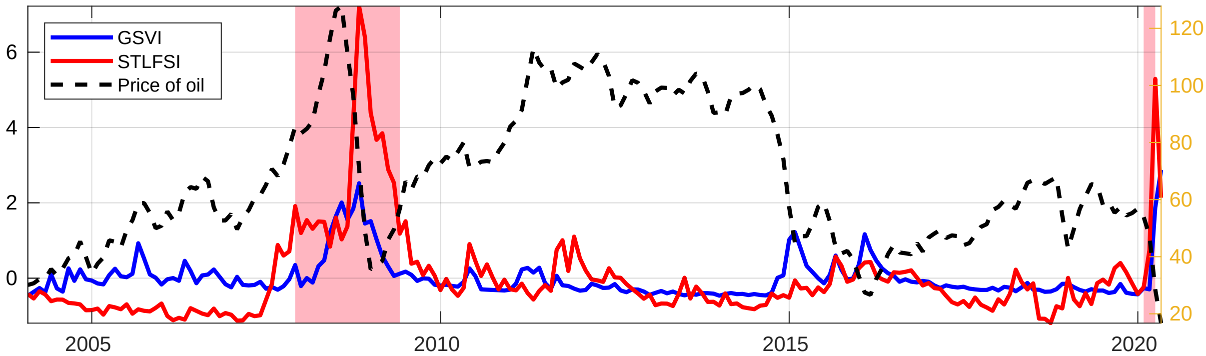
Over the last two decades, energy commodities have become a popular asset class for financial institutions and retail investors, similar to equities and bonds. Meanwhile, oil price fluctuations stimulated the academic and policy debate regarding the impact of fundamental shocks such as supply and demand shocks, inventory shocks, or macroeconomic and monetary shocks on the price of oil (see for instance Kilian, 2008, 2009; Leduc and Sill, 2004; Kilian and Murphy, 2014; Baumeister and Kilian, 2016). The price of crude oil may also be related to non-fundamental shocks, such as expectations or herd behaviour. Moreover, oil price shocks have evolved over time and have effects on the real economy through consumer and firm behaviour that are not constant (Blanchard and Riggi, 2013; Hamilton, 2013; Kang et al., 2015). This paper offers new insights on the relationship between oil prices and investors sentiment using a time-varying structural vector autoregression (SVAR) model.

With regard to the direction of causality between the crude oil prices and investors sentiment, the literature follows two main branches. On the one hand, under the efficient market hypothesis, the traditional asset pricing models assert that investors are rational and that information is instantaneously incorporated into asset prices. Departures from efficiency have mainly been investigated through some form of investors' reaction to new information (Barberis et al., 1998; Hong and Stein, 1999). In this respect, investors' attention or sentiment must play a significant role in asset prices and returns. A low level of investor attention may lead to small fluctuations in oil prices, while information that has received broad attention from investors will instantly be comprised in prices, resulting in large fluctuations (Vozlyublennaia, 2014). On the other hand, oil price fluctuations may lead to pessimistic expectations about future economic conditions and to a subsequent reduction in consumer and investor sentiment, which induces households and firms to curb their consumption and investment expenditures, respectively (Guntner and Linsbauer, 2018).

In this paper, we gauge investor sentiment for the oil market by employing a direct source of information provided by Google. Google Trends provides an index of the volume of Google queries by geographic location and category (for applications in financial markets see Da et al., 2011; Vlastakis and Markellos, 2012; Vozlyublennaia, 2014; Da et al., 2015). By inputting a keyword search term on Google Trends (for example, "oil price", "crude oil", "spot oil"), users can observe the actual flow of worldwide Internet searches for that particular keyword over time. This value is relative to the total number of search terms on Google in the corresponding time interval. To construct the oil index, we use 186 search terms related to oil prices and empirically identify most of their co-movement through a dynamic factor model analysis. The newly constructed index provides a measure for global information demand regarding the price of oil and accordingly represents a direct measure of investors' attention on the crude oil market.

The advantage of using international attention via Google data is that we can estimate people's active internet search and collective belief in different parts of the world at different frequencies (unlike surveys) and study their follow-up actions, as investors act and trade on their beliefs and move oil prices. Moreover, as investors take the initiative to search online information about the oil market, the Google search index possesses their own attention behaviour, so that bias is eliminated (Da et al., 2011). Sockin and Xiong (2015) argue that key industrial commodities, such as crude oil, often serve as important price signals regarding the strength of the global economy for market participants. Indeed, in Figure 1 we observe a strong co-movement of GSVI with the St. Louis Financial Stress Index (STLFSI), especially during the two most recent recession periods.[1] Therefore, Google searches can be considered a good proxy not only for aggregate investors' attention to oil markets but also for the overall outlook of the economy.

Figure 1: Oil prices, Google Search Volume Index and the Financial Stress Index over the period 2004-2020. Oil prices, measured as refiner's acquisition cost of crude oil, are plotted against the right axis. The shaded areas denote the two most recent recession periods.



Apart from investors' sentiment, economic policy uncertainty may exert a great influence on real economic activity. A number of studies show that the supply-side oil shocks are relatively unimportant for the macro-economy compared to demand-side oil price shocks (Kilian, 2009; Hamilton, 2009a; Lippi and Nobili, 2012; Baumeister and Peersman, 2013); this finding is confirmed by Antonakakis et al. (2014) and Kang and Ratti (2013) who investigate the relationship between structural oil price shocks and economic policy uncertainty. A widely used measure of uncertainty is the Economic Policy Uncertainty (EPU) index provided by Baker et al. (2016). The index is a weighted average of four uncertainty components, namely the news-based policy uncertainty index, the tax expiration's index, the CPI forecast disagreement measure and the federal/state/local purchases disagreement measure. In this respect, EPU assesses macroeconomic uncertainty by combining economic uncertainty related to

[1]The St. Louis Financial Stress Index (STLFSI) quantifies financial stress in the U.S. economy using 18 key indicators of financial market conditions - 7 interest rates, 6 yield spreads, and 5 other indicators.

economic policy making and public views. Baker et al. (2016) argue that a rise in EPU is related with an increase in price volatility and a decrease in investment.

Our empirical analysis builds on the structural vector autoregressive (SVAR) approach and extents the Kilian and Park (2009) model by including the investor's sentiment index for oil (GSVI). The extended model consists of five endogenous variables, the change in oil production, the global economic activity, the real oil prices, the stock market returns and the GSVI index. We further allow for time-variation in both the parameters and the standard deviation by employing a Bayesian estimation of the SVAR model. Our estimates show that the GSVI_{t-1} coefficient is negative in the stock returns equation for the entire sample period. Arguably, increases (decreases) in GSVI correspond with low (high) stock returns, a finding consistent with the sentiment induced mispricing hypothesis. Similar pattern is observed in the equation of real price of oil, but this relationship reverses after 2015 to positive. We also find that the standard deviation for all models increases considerably during the Global Financial Crisis in 2009, except the equation of stock market returns, where the standard deviation of the disturbance term reaches its highest value during the recent Covid-19 pandemic.

We further examine the importance of evolving parameters through the analysis of impulse response functions. We document that a positive shock in the GSVI variable yields a decrease in oil production which lasts only for two months. From this point on, oil production sharply increases and reaches pre-shock levels. Similar results are found in the responses of oil demand and oil prices. For the three aforementioned variables, the impulse responses are time invariant. Although, this is not the case with stock market: if the shock occurs before 2015, stock market returns decline after a couple of periods and if the shock occurs after 2015, stock market returns gradually increase. In both cases, the effect of the shock is both persistent and statistically significant.

Considering the role of economic policy uncertainty on firm level decisions and real economic activity, we subsequently estimate an alternative model by adding in the five-variable model the EPU index of Baker et al. (2016) as an indicator for economic policy uncertainty. Higher volatility in oil prices has been linked with increased uncertainty at firms. Elder and Serletis (2010) and Rahman and Serletis (2011) show that uncertainty about changes in the real oil prices has a negative effect on real economic activity. Moreover, Bird and Yeung (2012) show that during periods of high (low) market uncertainty investors tend to ignore good (bad) news. In this regard, the aim of this exercise is to uncover the time-varying role of economic policy uncertainty in the crude oil - investor sentiment relationship.

Sensitivity analysis shows that our baseline findings are robust. First, we employ as an alternative sentiment measure the Index of Consumer Expectations (ICE), produced by the University of Michigan, which summarizes a monthly aggregation of around 500 survey respondents' expectations regarding the future outlook for their own personal finances and for the overall U.S. economy (see, e.g., Carroll et al., 1994; Matsusaka and Sbordone, 1995; Lagerborg et al., 2020). Second, we use specifications that consider the role of oil inventories to show that the results are not driven by an "identification problem" (see for example Hamilton, 2009a,b; Kilian and Murphy, 2014; Baumeister and Hamilton, 2019).

Our paper contributes to the literature in several aspects. A great deal of literature has focused on verifying the impact of investor sentiment on crude oil returns by applying a limited number of search terms. Our paper constructs a market specific investor sentiment index for crude oil markets based on a much wider set of attention terms and using the dynamic factor model (DFM) framework. This framework allows us to consider dynamics both in the factors and the idiosyncratic component, as well as heteroscedasticity in the idiosyncratic variance. Based on Bayesian estimation, this paper also studies the changing effects of the degree of investor sentiment in crude oil markets. Extending the model of Kilian and Park (2009) with the oil GSVI index, we find that information demand responds temporarily to oil supply and demand shocks and shocks in stock market returns. Also, by allowing for time-variation we find that these responses are time dependent (for example, a shock in oil prices after 2015 has a greater impact on GSVI). We further examine the importance of the changes of parameters of impulse response functions, where the sources of time variation are both the coefficients and the variance covariance matrix of the innovations. Doing so, we may identify changes in the impulse responses, mainly in magnitude but for some cases also in the sign of the response. Furthermore, to the best of our knowledge, this is the first paper that explores the role of economic policy uncertainty in the linkage between investor attention and crude oil market through a time-varying framework. Our findings show that the effect of policy uncertainty is time-varying and becomes significant and positive only during the period preceding the 2009 banking crises.

The rest of the paper is organised as follows: Section 2 describes the methodology. Section 3 describes the data. Section 4 presents the main findings of the study. Section 5 investigates the robustness of our results. The last section concludes.

2 Econometric methodology

Initially we construct the index for investors' sentiment based on internet searches for oil-related terms. The construction of the sentiment index is described in section 2.1. Once the new index is created, we estimate a TVP-VAR model with five variables: change in oil production, global real economic activity, real oil prices, stock market returns and investors' sentiment. We examine the effect of investors' sentiment on oil price fundamentals and stock market returns by examining the evolution of the estimated coefficients over time. In addition, we employ impulse response function analysis to examine the bilateral relationship between investors' sentiment and oil supply, oil demand, oil prices and stock market.

The estimation of the TVP-VAR model is described in section 2.2. In the last part of the analysis we extend the main model by including a variable for economic policy uncertainty. In this case, we focus on the effect of a shock from and to the policy uncertainty variable.

2.1 Construction of the Google Search Volume Index

We start by constructing a list of sentiment-reveal search terms that are related to the crude oil market. We initiate the analysis with primary keywords such as "oil", "oil price", "crude oil","oil spot", "oil stock" etc.[2] More specifically, we have considered each search term in Google Trends search engine and we have kept the ten "top searches" related to each term. This generates approximately 300 related keywords for oil. Next, we remove duplicates, terms with no economic meaning and terms with inadequate data (query series with zero search volume). This leaves us with the final 186 oil-related search terms.

To construct the Google Search Volume Index (GSVI) we download from Google Trends the monthly SVI for 186 search terms related to oil, from January 2004 to May 2020 and employ the dynamic factor model (DFM) framework.[3] DFM is appropriate for handling large data sets without much loss of information.[4] In addition, DFM allows for more flexible interpretation compared to principal components analysis and provides a better fit to the data (Sargent and Sims, 1977; Stock and Watson, 2016).

The DFM is written in the state space representation as:

$$x_t = \sum_{i=0}^{s} \Lambda_i f_{t-i} + e_t \tag{1}$$

$$f_t = \sum_{t=1}^{p} A_i f_{t-i} + u_t \tag{2}$$

where x_t is an $N \times 1$ vector of observable variables, Λ_i are $N \times q$ matrices of factor loadings, f_t are $q \times 1$ vectors of unobserved common dynamic factors which summarise the cross-covariance properties of x_t. The idiosyncratic component e_t, are $N \times 1$ stationary processes uncorrelated with f_t. The elements of e_t are weakly correlated cross-sectionally and/or serially. The VAR(p) model in equation 2, with $A_1 \ldots, A_p$ the matrices of autoregressive coefficients, approximates the dynamics of the common factors. The two error vectors are assumed to be independent.

The DFM described by equations 1 and 2 can be transformed into a static state space representation.

[2]Table A1 reports all keywords used for the construction of the index.

[3]All time-series are stationary considering two unit root tests (the Augmented Dickey-Fuller test and the Phillips-Perron test) and are properly scaled to the highest value within the whole sampling period.

[4]For example, principal components analysis can lead to information loss if the number of components is not specified correctly.

We define the state vector as $F_t = (f_t', \ldots, f_{t-k+1}')'$ where $k = \max\{s+1, p\}$. If $k > p$ then $A_{p+1} = \cdots = A_k = 0$ in the companion matrix:

$$\mathbf{A} = \begin{bmatrix} A_1 & \ldots & A_p & \ldots & A_k \\ I_q & 0_q & . & . & 0_q \\ 0_q & \ddots & \ddots & . & \vdots \\ \vdots & \ddots & \ddots & . & 0_q \\ 0_q & . & 0_q & I_q & 0_q \end{bmatrix}$$

if $s+1 < k$, we set $\Lambda_{s+1} = \cdots = \Lambda_k = 0$ in the loadings matrix $\Lambda = (\Lambda_0\ \Lambda_1 \ldots \Lambda_k)$. The static factor form of the state-space model is given by the following equations:

$$x_t = \Lambda F_t + e_t$$
$$F_t = \mathbf{A} F_{t-1} + u_t^*$$

with $u_t^* = (u_t', 0_{1\times q}, \ldots, 0_{1\times q})'$. The number of estimated static factors $K = q(s+1)$ is determined using the information criterion methods provided by Bai and Ng (2002). The number of dynamic factors given a preselected number of static factors is computed using the four criteria Bai and Ng (2007).

The factors are estimated using Quasi ML - EM estimator developed by Doz et al. (2012) by iterating the two-step estimator of Doz et al. (2011). First an 'extended' state space model is defined, with one more lag than the representation (1),

$$x_t = \sum_{i=0}^{s} \Lambda_i f_{t-i} + \mathbf{0}_{N\times q} f_{t-s-1} + e_t$$

where $\mathbf{0}_{N\times q}$ indicates an $N \times q$ null matrix. The static representation of this system then becomes:

$$x_t = \grave{\Lambda} G_t + e_t$$
$$G_t = \grave{\mathbf{A}} G_{t-1} + \grave{u}_t^*$$

where $G_t = (F_t', f_{t-k}')' = (f_t', \ldots, f_{t-k+1}', f_{t-k}')'$ and the system matrices are redefined accordingly. This modification of the state space representation is necessary to make it possible to compute rapidly and efficiently all the quantities that are required in the M-step. Therefore, given a preliminary estimate of $\hat{\theta}$, the 'extended' system matrices are computed and the Kalman smoother is run, so as to get an unbiased predictor of the factors $\hat{G}_t$ with the associated covariance matrices. This is the E-step. Subsequently, the estimated factors are used to compute the sufficient statistics:

$$\hat{F}_t = E(F_t|\mathbf{X})$$
$$\hat{P}_{0,t} = E(F_t F_t'|\mathbf{X})$$
$$\hat{P}_{1,t} = E(F_t F_{t-1}'|\mathbf{X})$$

where $\mathbf{X} = x_1, \ldots, x_T$, which are used in the M-step to compute (conditional) ML estimates of θ. Note that, given the covariance matrix of G_t, the calculation of $\hat{P}_{0,t}$ and $\hat{P}_{1,t}$ simply amounts to selecting appropriate submatrices. With the new estimate of θ in hand, the process is repeated until convergence. The first $q-$ element subvector $\hat{f}$ of $\hat{F}_t$, the ML estimate of the dynamic factor, is used as the GSVI.

To control convergence of the EM algorithm we employ a parameter distance criterion that adds the absolute deviation across all estimated parameters $\hat{\theta}$ between two consecutive steps. More formally, we calculate:

$$c_P^j = \frac{\sum_{i=1}^{h} |\hat{\theta}_I^{(j)} - \hat{\theta}_I^{(j-1)}|}{h} \tag{3}$$

where h is the number of elements in $\hat{\theta}$. We stop EM process when $c_M^P < 10^{-2}$.

Figure 1 plots the constructed index over the full sampling period, 2004M01-2020M04.

2.2 Estimation of the TVP-VAR model

Based on Primiceri (2005), Del Negro and Primiceri (2015) and Koop and Korobilis (2009) we consider the following time-varying parameter VAR model:[5]

$$y_t = c_t + \sum_{i=1}^{p} B_{i,t} y_{t-i} + u_t, \quad t = 1, \ldots, T \tag{4}$$

where y_t is an $n \times 1$ vector of endogenous variables, c_t is a $n \times 1$ vector of time-varying coefficients, $B_{i,t}$, $i = 1, \ldots, P$ are $n \times n$ matrices of time-varying coefficients and u_t are heteroscedastic shocks with variance-covariance matrix Ω_t. The lag order p is determined through the Schwarz information criterion (SIC). We can rewrite (4) as:

[5] In this section we drop any notation used in the section 2.1.

$$
\begin{aligned}
y_t &= X_t^{'} B_t + A_t^{-1} \Sigma_t \epsilon_t \\
X^{'} &= I_n \otimes [1, y_{t-1}, \ldots, y_{t-p}]
\end{aligned} \tag{5}
$$

where $B_t = [c_t, \text{vec}(B_{1,t}), \ldots, \text{vec}(B_{p,t})]$ and $\otimes$ denotes the Kronecker product and vec() the columnwise vectorisation of a matrix. A_t is a lower unitriangular matrix and Σ_t is a diagonal matrix such that, $A_t \Omega_t A_t^{'} = \Sigma_t \Sigma_t^{'}$. It follows that $\text{Var}(\epsilon_t) = I_n$. Cogley and Sargent (2002) consider similar decompositions of the variance-covariance matrix of a time-varying VAR model but with a time invariant matrix A_t. If α_t is the vector of unrestricted elements of A_t and σ_t the vector of diagonal elements of Σ_t, then the state equations are:

$$
B_{t+1} = B_t + v_{t+1}, \quad v_t \sim N(0, Q) \tag{6}
$$

$$
\alpha_{t+1} = \alpha_t + \zeta_{t+1}, \quad \zeta_t \sim N(0, S) \tag{7}
$$

$$
\log \sigma_{t+1} = \log \sigma_t + \eta_{t+1}, \quad \eta_t \sim N(0, W) \tag{8}
$$

The cardinalities of B_t, a_t and σ_t are $n_B = n + pn^2$, $n_A = \frac{n(n+1)}{2}$ and $n_\sigma = n$. All errors in the model are assumed to be jointly normally distributed with

$$
\text{Var}\left(\begin{bmatrix} \epsilon_t \\ v_t \\ \zeta_t \\ \eta_t \end{bmatrix} \right) = \begin{bmatrix} I_n & 0 & 0 & 0 \\ 0 & Q & 0 & 0 \\ 0 & 0 & S & 0 \\ 0 & 0 & 0 & W \end{bmatrix}
$$

As in Cogley and Sargent (2005) and Primiceri (2005), the prior distributions (initial conditions at $t = 0$) are specified and updated using a training sample. The training sample is used to estimate a time invariant VAR with Ordinary Least Squares (OLS). For the parameters B_0, A_0 and $\log \sigma_0$, the Normal distribution is assumed. The mean and variance of the parameters B_0 and A_0 are the mean and four times the variance of the respective estimated parameter of the OLS-VAR. In the case of $\log \sigma_0$ the mean of the prior distribution is chosen to be the logarithm of the OLS point estimates of the standard errors of the OLS-VAR model and the the variance covariance matrix is assumed to be the identity matrix. For the priors on error co-variances Q, S and W the hyperparameters are set to $k_Q = 0.01$, $k_S = 0.1$ and $k_W = 0.01$. Summarising, the priors take the form:

$$
\begin{aligned}
B_0 &\sim N(\hat{B}_{OLS}, 4\text{Var}(\hat{B}_{OLS})) \\
A_0 &\sim N(\hat{A}_{OLS}, 4\text{Var}(\hat{A}_{OLS})) \\
\log \sigma &\sim N(\log \hat{\sigma}_{OLS}, I_n) \\
Q^{-1} &\sim W(1 + n_B, ((k_q^2)(1 + n_B)\text{Var}(\hat{B}_{OLS}))^{-1}) \\
S^{-1} &\sim W(1 + n_A, ((k_q^2)(1 + n_A)\text{Var}(\hat{A}_{OLS}))^{-1}) \\
W^{-1} &\sim W(1 + n_\sigma, ((k_q^2)(1 + n_\sigma)I_n)^{-1})
\end{aligned}
$$

Let M_t be a generic matrix. We define $M^\tau = [m_1', \ldots, m_\tau']'$, where $m_t = vec(M_t)$. Our aim is to estimate the posterior distributions of the unobservable states B^T, A^T and Σ^T and the hyperparameters of the variance covariance matrix V. The Bayesian estimation requires a Markov chain Monte Carlo (MCMC) algorithm. To exploit the blocking structure of the unknowns, a Gibbs sampling consisting of four steps is used. In each step B^T (the time-varying coefficients), A^T (the simultaneous relations), Σ^T (the stochastic volatility) and V (the hyperparameters) are drawn respectively, conditional on the observable data and the other parameters.

The state space model described in 5 and 6 is linear and Gaussian (conditional on A^T and Σ^T). This implies that B^T, A^T is a product of Gaussian densities and can be drawn using a Gibbs sampler algorithm.[6] Here, we rely on the algorithm of Carter and Kohn (1994).[7] Since Σ^T is not a Gaussian product, we rely on the method of Kim et al. (1998) to transform it to a linear, Gaussian form. As a result, it can be drawn using the algorithm of Carter and Kohn (1994). Finally, we sample V by sampling Q, W and S independently from an inverse-Wishart distribution.

Once the reduced-form VAR model (4) is estimated, we construct the following structural VAR:

$$
y_t = X_t' B_t + \Xi \epsilon_t
$$

where Ξ_t contains the necessary restrictions for structural identification and $\Xi_t \Xi_t' = \Omega_t$. Since the identification is based on a lower triangular scheme, it suffices to set $\Xi = A_t^{-1} \Sigma_t$.

The results are based on posterior sample of 50000 draws. The first 25000 draws are considered a burnin sample and are discarded. To assess the validity of our results we considered two robustness checks. In the first one, we increase the number of both burnin and posterior draws to 50000. The second check aims to take into account any possible autocorrelation among draws, (Korobilis, 2017). To this

[6] A^T is drawn conditional on B^T and Σ^T

[7] Alternatively, one could employ the algorithm of Frühwirth-Schnatter (1994).

end, we build the posterior sample by keeping only every 10$^{\text{th}}$ draw. Both robustness checks provide similar results with the ones reported in the paper. These results are available upon request.

3 Data

The study utilises monthly time-series data over 2004M01-2020M04. Data for the oil production and prices of oil are obtained from the U.S Department of Energy. World crude oil production is measured in millions of barrels per day averaged monthly. Following Kilian (2009) we use the percent change in the oil supply by calculating the log differences and multiplying with 100 (denoted $\Delta prod_t$). To obtain the real price of oil (denoted rpo_t), we divide the refiner's acquisition cost of crude oil by the U.S. CPI (obtained from the Bureau of Labor Statistics).[8] The index of real aggregate demand, constructed by Kilian (2019) is used as an indicator of global real economic activity(denoted rea_t). The logarithm of closing prices of S&P500, deflated by the U.S. CPI is used as the stock market variable (denoted ret_t). As a proxy for economic policy uncertainty, we employ the index constructed by Baker et al. (2016) which is a weighted average of four uncertainty components (denoted epu_t). Finally, the sentiment index, GSVI, is constructed as described in subsection 2.1. SVI data used for the construction of the GSVI is available from 2004M01. The availability of SVI determines the sample size in our analysis.

In our analysis, we also perform robustness checks regarding the first main (five-variables) model. First, we replace GSVI with Michigan's University Index of Consumer Sentiment (ICS) (see section 5). ICS is constructed using the responses to five different questions that are part of a broader survey of consumer attitudes and expectations. Data for the ICS are available on monthly basis starting from January 1978 from the website of Michigan's University. Second, we examine how the results are affected when we take oil inventories into account. To construct the variable of crude oil inventories, we multiply the U.S. inventories by the ratio of OECD inventories (see also Hamilton, 2009a; Kilian and Murphy, 2014; Baumeister and Hamilton, 2019). We obtain data for the U.S. crude oil stocks (in millions of barrels) and for the OECD inventories of crude petroleum and petroleum products from the Energy Information Administration and the OECD databases, respectively. Table A2, in the Appendix, reports the summary statistics for all variables used in the analysis.

[8] Kuck and Schweikert (2017) and Galay (2019) argue that world oil markets maintain a long-run equilibrium.

[8] The index is an updated and corrected version of the Kilian (2009) index. See also Kilian and Zhou (2018) and Hamilton (2021).

4 Results

4.1 The effect of GSVI on the price of oil

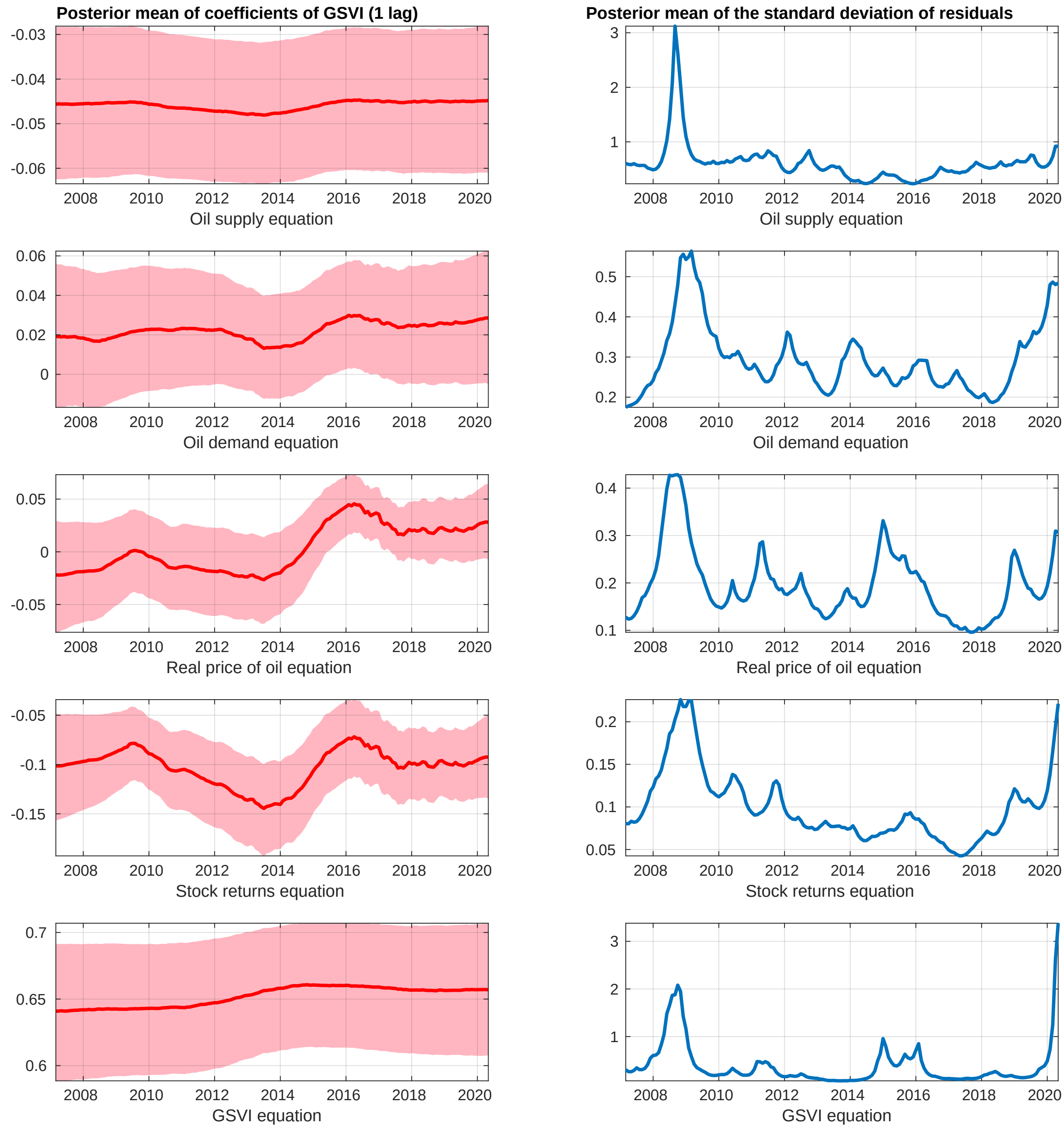
Figure 2 shows the evolution of the estimated parameters over time (we consider two lags in the model based on the SIC).[9] The left column of Figure 2 presents the posterior means of the coefficients of the first lag of the GSVI variable (1 lag) in the five equations of the model: the oil production, oil demand, real price of oil, stock market returns and information demand equation. The shaded area in each graph reports the 68% credible interval. In the oil supply equation, the coefficients of lagged GSVI are negative, statistically significant and remain stable over the entire sample period. In the equation of oil demand, the coefficients of GSVI are positive and statistically insignificant for most of the sampling period. For a brief period, during 2016, the coefficient becomes statistically significant. There does seem to be a movement in the coefficients of the variable of information demand in the equation of real price of oil. The posterior coefficient increases over the first 5 years of the sample and then gradually declines until 2014. After 2014 it increases for almost two years and then remains relatively constant. In addition, the posterior mean of coefficients is negative until 2015 and it becomes positive after that period. We observe a similar pattern in the evolution over time of coefficients of GSVI in the stock market returns equation. However, the values of the coefficients are negative over the whole sample.

The effect of time is also evident on the standard deviation of the errors which justifies the selection of a model that accounts for stochastic volatility. The right column in Figure 2 presents the posterior means of the standard deviation of residuals in the five equations of the VAR model. In all cases, there is a steep increase in the standard deviation during 2009, caused by the Global Financial Crisis. Furthermore, there are cases in the sample where the standard deviations of the shocks increase for a short period. Finally, it is worth noticing the effect of the recent pandemic in the volatility of the error terms. Especially in the case of stock market returns where the standard deviation of the disturbance term is currently increasing and tends to exceed the maximum value which reached during 2009. The rate of increase is also higher than in 2009.

We now turn to the impulse response function analysis. The identification scheme builds on the work of Kilian (2009) and Kilian and Park (2009). We consider the following relationship between the reduced form errors u_t and structural errors ϵ_t:

[9] The selection of lag order is consistent with similar studies in the literature (Apergis and Miller, 2009; Degiannakis et al., 2013; Kang et al., 2015).

Figure 2: Posterior means of coefficients of lagged GSVI and posterior means of standard deviations of the residuals in the mean equations.



$$\begin{bmatrix} u_t^{\Delta prod} \\ u_t^{rea} \\ u_t^{rpo} \\ u_t^{ret} \\ u_t^{gsvi} \end{bmatrix} = \begin{bmatrix} \xi_{11} & 0 & 0 & 0 & 0 \\ \xi_{21} & \xi_{22} & 0 & 0 & 0 \\ \xi_{31} & \xi_{32} & \xi_{33} & 0 & 0 \\ \xi_{41} & \xi_{42} & \xi_{43} & \xi_{44} & 0 \\ \xi_{51} & \xi_{52} & \xi_{53} & \xi_{54} & \xi_{55} \end{bmatrix} \begin{bmatrix} \epsilon_t^{\Delta prod} \\ \epsilon_t^{rea} \\ \epsilon_t^{rpo} \\ \epsilon_t^{ret} \\ \epsilon_t^{gsvi} \end{bmatrix}$$

where we assume that stock market and information demand react contemporaneously to all supply and demand shocks and we treat shocks in oil prices as predetermined with the respect to the economy (Lee and Ni, 2002).

Our aim is to examine the bilateral relationship between GSVI and the price of oil. To this end, we consider the responses of the variables to shocks in the GSVI and the response of GSVI to shocks to the rest of the variables. The top row in Figure 3 presents the impulse responses of the variables to a shock in information demand. A positive shock in the GSVI variable yields a decrease in oil production which lasts only for two months. After that, oil production sharply increases and reaches pre-shock levels. The shock also causes a rise in aggregate demand which in turn yields a rise in the price of oil. Specifically, we observe a positive co-movement in aggregate demand and oil prices with the impact on the price of oil being greater than on aggregate demand. This result, combined with the delayed increase in oil production indicates that an unexpected increase in information demand acts as a mechanism for a shock in real economic activity, leading eventually in a positive co-movement of oil production and the price of oil. Finally, a sudden increase in information demand does not have an immediate effect on stock market. In addition, the effect is different, depending on the timing the shock in GSVI occurs: a shock in GSVI before 2015 causes stock market returns to start to fall four months after the shock. On the contrary, a similar shock after 2015 causes a rise in stock market returns after four months. In both cases, the effect of the shock is long-lasting and the upward/downward trends continues for more than two years.

The previous results suggest that unexpected changes in GSVI affect the behaviour of oil dynamics and stock market returns. However, oil demand and supply shocks may affect household disposable income available for other expenditures through energy prices. Since prior evidence suggests that oil price shocks are mainly transmitted through the demand side, we investigate the effect of structural oil supply and demand shocks on GSVI—a barometer of private households', retail investors and non-institutional traders perception of uncertainty and future economic conditions. The bottom row of Figure 3 shows the impact of different shocks on the GSVI index. All shocks affect investors' sentiment on impact, reflecting the rapid spread of news worldwide. Furthermore, the sign of the response is affected by the time the shock occurs. For example, disruptions of the physical supply, which raise the real crude oil price, have a limited impact on the GSVI until 2010, but this effect is becoming stronger as we move to 2015 onward. The diminishing role for oil supply shocks suggests that households and retail investors expect that these shocks are short-lived, given that reduced production in one country is easily offset by other oil producers. On the demand side, both aggregate demand shocks associated with the global business cycle and oil-specific demand shocks significantly depress GSVI during the period prior to 2015. This impact turns to positive in subsequent years but is short-lived. In line with the results in Kilian (2009) and Guntner and Linsbauer (2018), positive aggregate demand shocks cause an increase of optimism among retail investors, followed by a significant reduction in the GSVI over the next 18 months. Other oil demand shocks such as an oil-specific demand shock, have a strong persistent positive effect on the GSVI over the recent years. The impact of a shock in stock returns on GSVI is positive and temporary over the last two years and during the global financial crisis period.

For illustrative purposes we also consider the impulses responses from a model with constant parameters. We estimate the impulse response functions using both an OLS-SVAR estimator and local projections (LP) by Jordá (2005). Following the related literature, in all time invariant models we use 24 lags.

The results are presented in Figure 4. In most cases, SVAR and LP produce qualitatively similar results.[10] LP produce narrower confidence intervals and as a result, suggest statistical significance of the responses more often compared to SVAR estimates.[11] The responses of aggregate demand and oil prices remain increasing, similar to the TVP-VAR model. The information demand shock leads to a decrease in stock market returns which is statistically significant only when we consider LP. Finally, the shock has no impact on crude oil production. The response of the GSVI to shocks in the rest of the variables is a temporary increase.

[10] This confirms the findings of Plagborg-Moller and Wolf (2021b)

[11] The LP are used as a less biased, robust to model specification estimator in contrast to the biased but efficient SVAR (Plagborg-Moller and Wolf, 2021a).

Figure 3: Impulse responses to an investors' sentiment shock (top) and impulses responses of GSVI to alternative shocks (bottom).

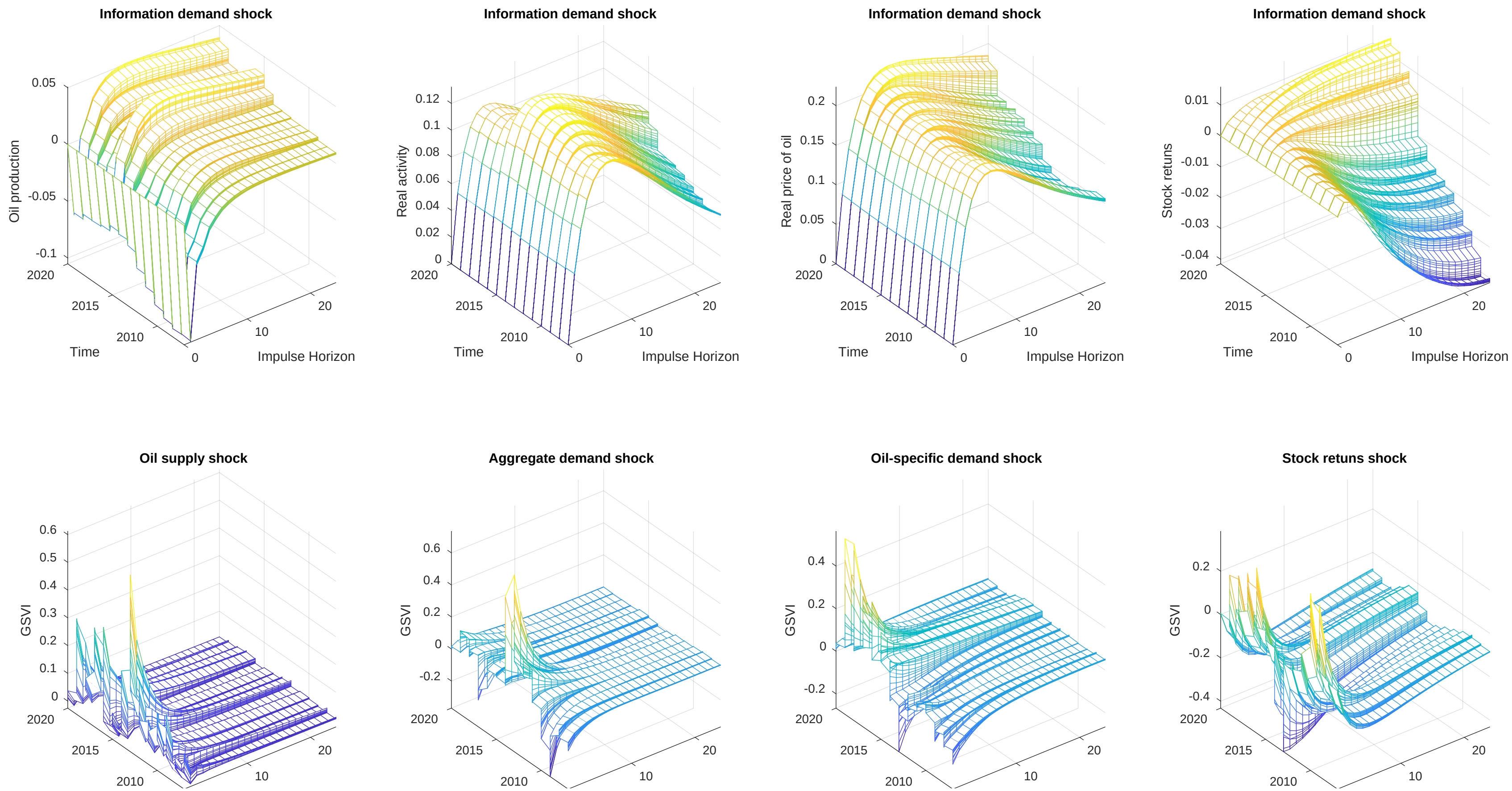
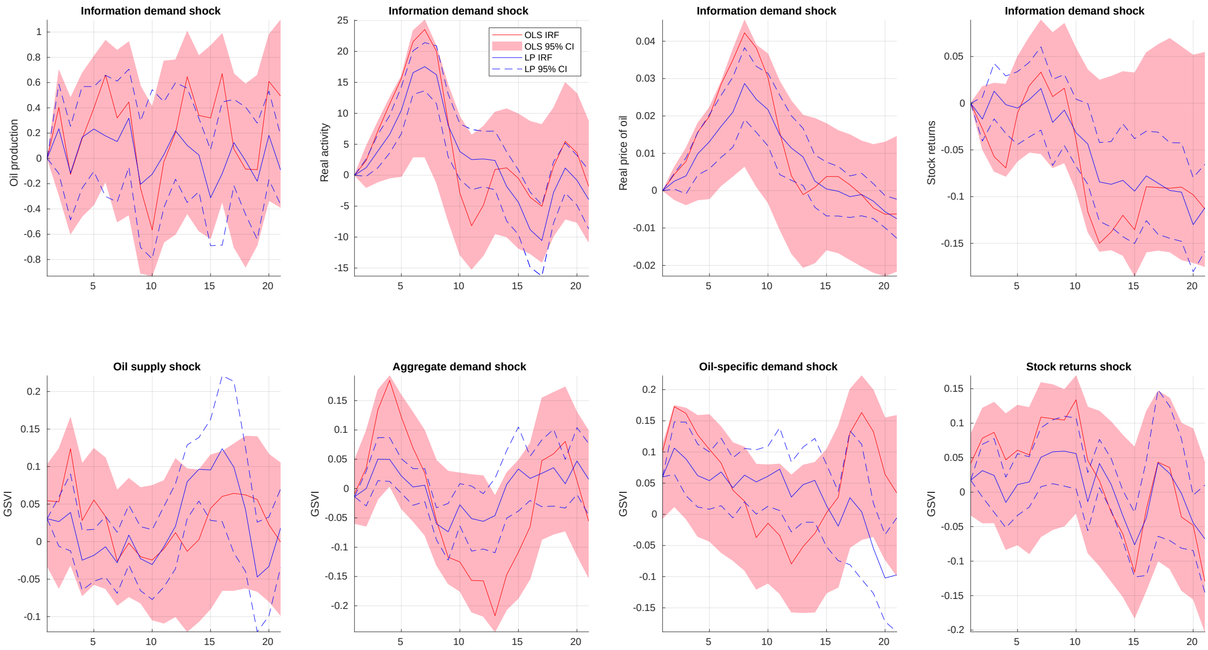


Figure 4: Impulse responses using time invariant SVAR and LP estimators. The first row plots the responses of the variables to an investors' sentiment shock. The second row plots the responses of GSVI to alternative shocks.



4.2 The impact of policy uncertainty on oil prices

In this section, we augment the baseline model by adding the EPU index. This model is based on the work of Kang and Ratti (2013) but differs in two respects. First, our model includes a measure of investors' sentiment and second, we consider a time-varying SVAR approach. By the same reasoning as in the previous model, the relationship between the reduced form and the structural errors is $u_t = \Xi \epsilon_t$ where $u_t^{'} = (u_t^{\Delta prod}, u_t^{rea}, u_t^{rpo}, u_t^{ret}, u_t^{gsvi})$, $\epsilon_t^{'} = (\epsilon_t^{\Delta prod}, \epsilon_t^{rea}, \epsilon_t^{rpo}, \epsilon_t^{ret}, \epsilon_t^{gsvi})$ and Ξ is 6×6 lower triangular matrix.

Figure 5 presents the impulse responses of the variables to a shock in the EPU index. The shock has an impact on oil production only in the short-run. Over the period 2007-2010 oil production increases after the shock while the after 2010, the shock causes a decrease. The shock has an negative impact on real economic activity if it occurs early in the sampling period. Similar response is observed for the oil prices. The response of stock market returns on the EPU index depends on time. From 2007 to 2008, an unanticipated increase in EPU affects stock market returns negatively in the short-run (the first ten periods). On the contrary, a shock after 2008 causes a slight increase in stock returns. Finally, a shock in EPU has a marginal impact on GSVI which lasts only for a few periods.

Figure 6 plots the responses of the EPU index to a shock in the rest of the variables. An unexpected disruption in oil supply has no significant effect on the EPU index. On the contrary, a positive shock in aggregate demand leads to an increase in policy uncertainty. After the initial rise, EPU gradually decreases. However, over the recent years the effect of the shock persists until the end the impulse horizon. Shocks in oil prices and stock market returns have a similar effect on the EPU index. Both shocks yield a drop in policy uncertainty over the first two periods (months) which gradually returns to pre-shock levels. GSVI is affected from a shock in EPU only on impact. With the exception of a brief period from 2007 to 2008 when a shock causes an increase in information demand, a sudden rise in EPU has a weak negative effect on GSVI.

We now present the findings from a time invariant model. The top row of Figure 7 plots the impulse responses of the variables to a positive shock in EPU. Both the SVAR and the LP estimators produce similar results regarding the direction of each response, however, only the LP provide statistically significant estimates. An increase in EPU has a significant negative effect on up to fifteen months after the shock. Based on the SVAR estimates, the shock has a negative but insignificant impact on the price of oil. The shock also affects negatively stock market returns. When we use LP, we observe a substantial decrease in stock market returns that lasts up to thirteen months after the shock. Finally, the sentiment index decreases only for a couple of periods after the shock.

The second row of Figure 7 presents the responses of EPU. A positive aggregate demand shock has a negative effect on EPU that is statistically significant up to two months and from thirteenth to fifteenth month. Similarly, an oil-specific demand shock decreases EPU only on impact. A shock in stock market returns leads to an immediate decrease in EPU that lasts up to three periods after the shock. After that period, EPU rises to pre-shock levels for five months and then declines again for five more months. Finally, an increase in information demand does not affect EPU.

Figure 5: Impulse responses to an EPU shock.

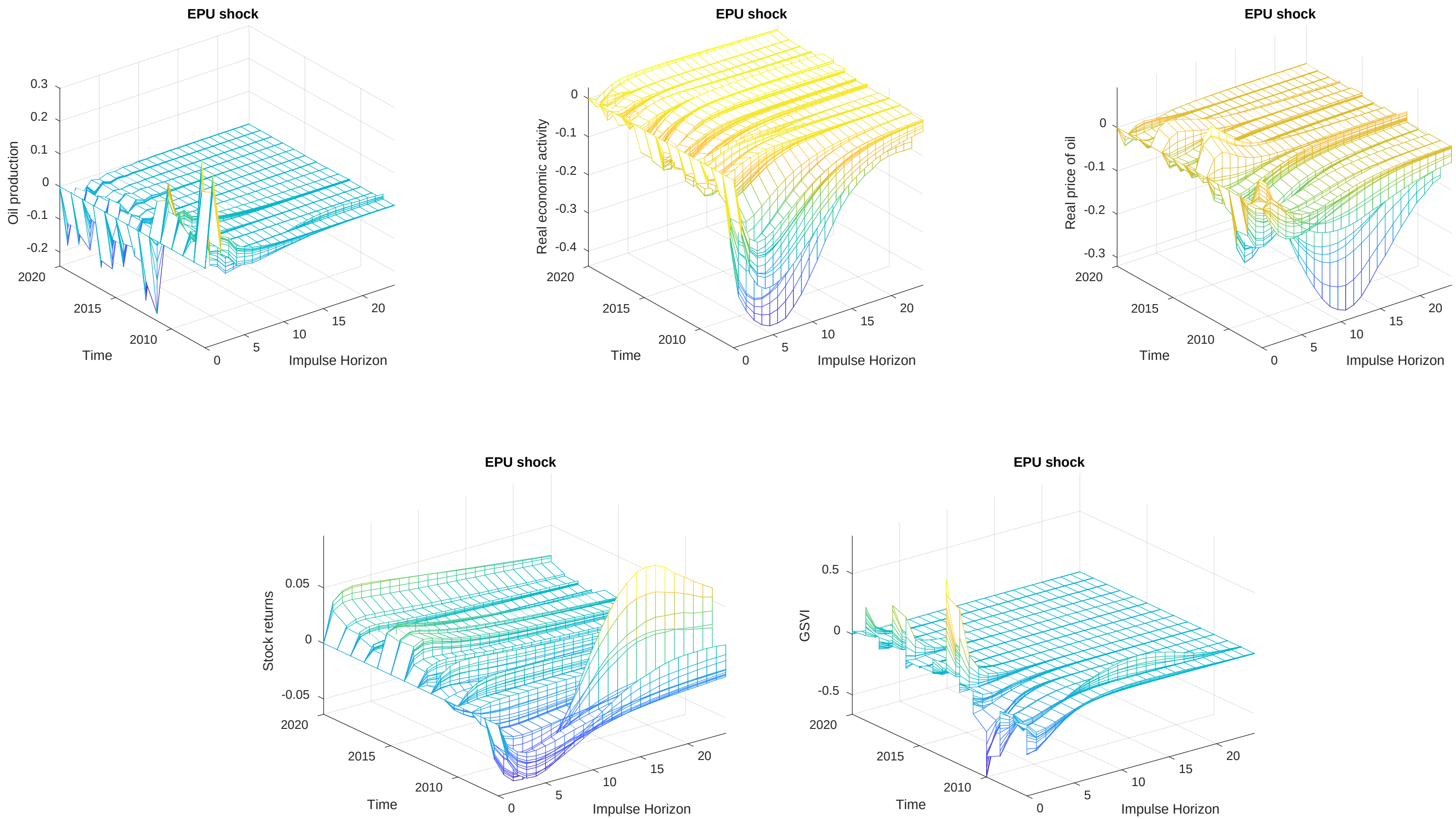


Figure 6: Impulse responses of the EPU index to alternative shocks.

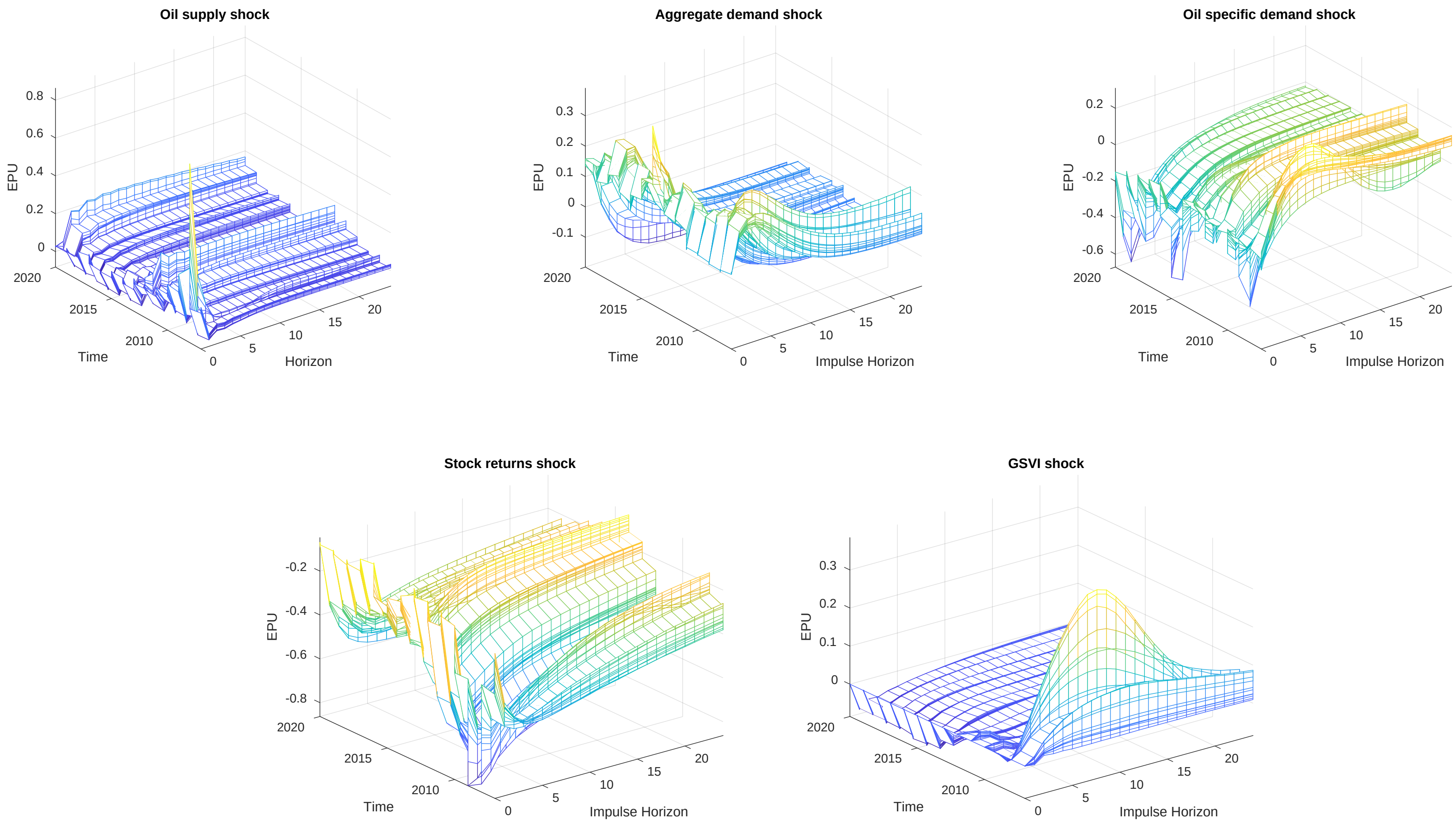
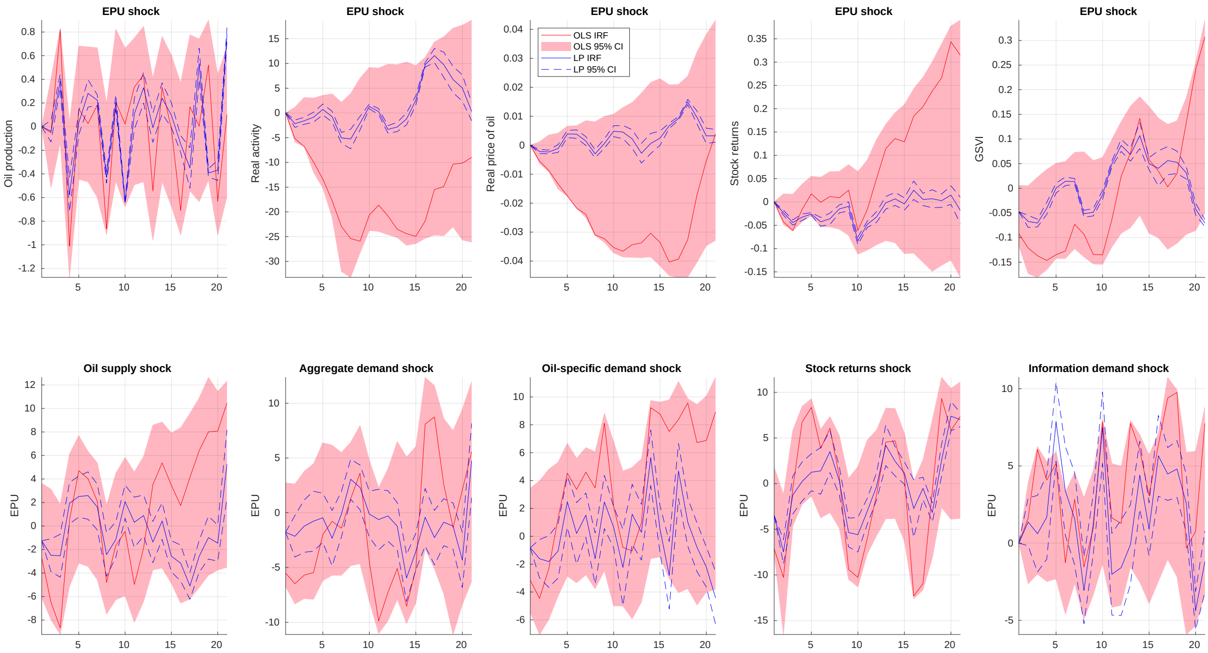


Figure 7: Impulse responses using time invariant SVAR and LP estimators. The first row plots the responses of the variables to an EPU shock. The second row plots the responses of EPU index to alternative shocks.



5 Robustness checks

5.1 Replacing GSVI with ICS

In this section we perform a robustness check regarding the model discussed in section 4.1 (the five-variable model). To examine the robustness of the results we use an alternative sentiment index which is often used in the literature, the Index of Consumer Sentiment (ICS) of Michigan's University (for applications, see Carroll et al., 1994; Matsusaka and Sbordone, 1995; Lagerborg et al., 2020). The index is available from 1978, however, for comparability purposes we estimate the model over the period 2004M01-2020M04.[12] Furthermore, we use the same number of lags, 2, as in the main model.

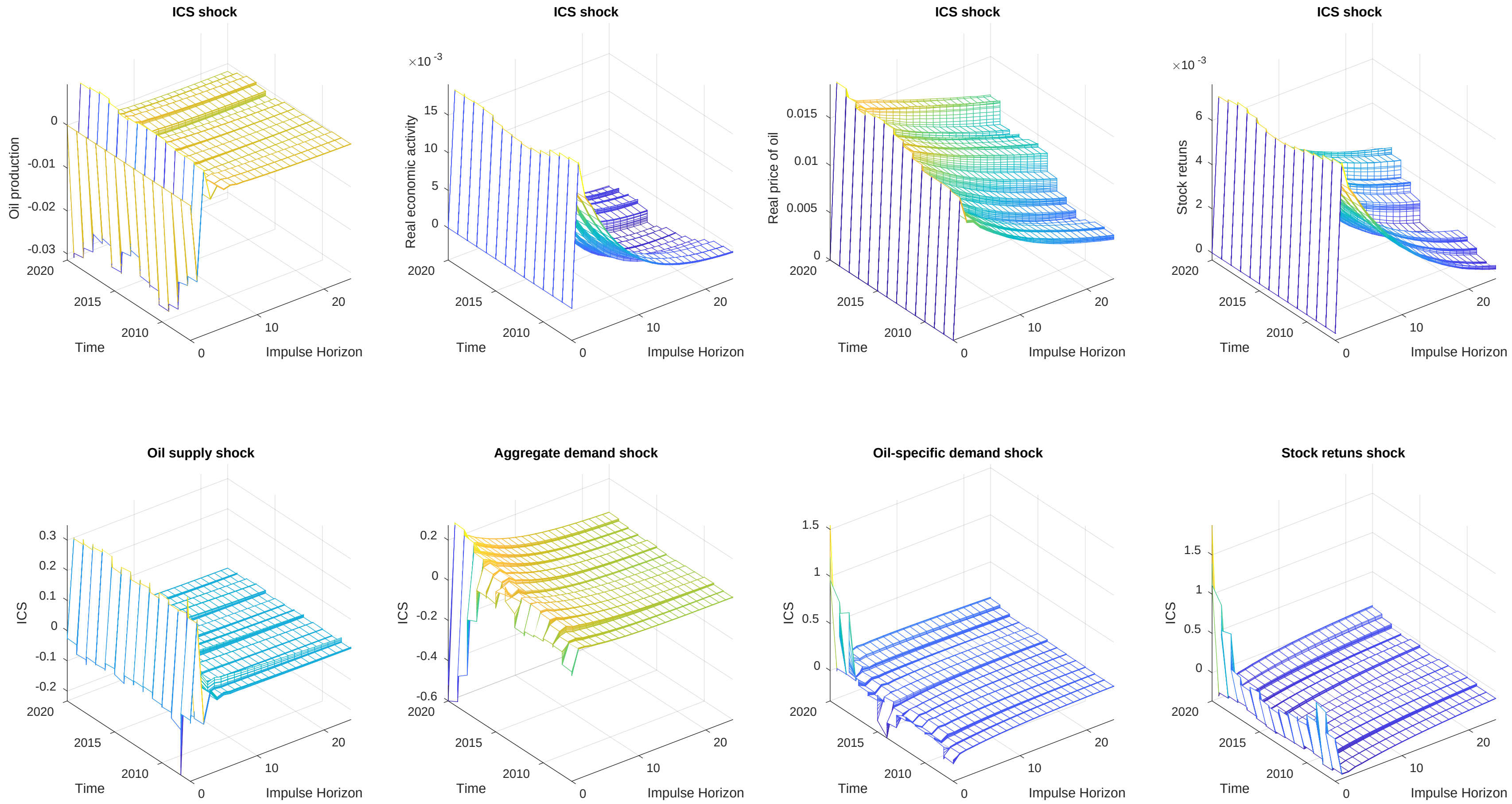
The first row in Figure 8 shows the impact of a positive shock in the ICS. Oil supply responds with an initial decrease which is reversed over the next month. After that initial response, oil production remains unresponsive. Real economic activity index, oil prices and stock returns are affected positively by the shock. The effect is more persistent in the case of oil prices and stock market returns. Overall, the responses of the first three variables are similar to the ones in the main model. However, the magnitude of the response is now smaller and independent of the time of the occurrence of the shock. Even in

[12]The results regarding the full sample are available from the authors upon request.

periods of known events (i.e. Global Financial Crisis), we observe no change in the response of the variables to alternative shocks. The latter lead us to conclude that the ICS can not capture the effect of time evolution. This explains the difference in the responses of stock market returns to shocks in GSVI and ICS over the first half of the sample.

The second row of Figure 8 presents the responses of ICS to shocks in the rest of the variables. Similar to GSVI, ICS is affected only for a couple of months from these shocks. While the response of ICS to each shock is similar over time, there is an increase in the magnitude of the response over the period 2018-2020 (oil production is an exception). Furthermore, while real economic activity increases after a shock from 2007 to 2018, after 2018 the shock causes a sharp decline in real economic activity.

Figure 8: Impulse responses using an alternative measure of investors' sentiment (GSVI is replaced by the ICS). The figure shows the impulse responses to shocks to and from the ICS.

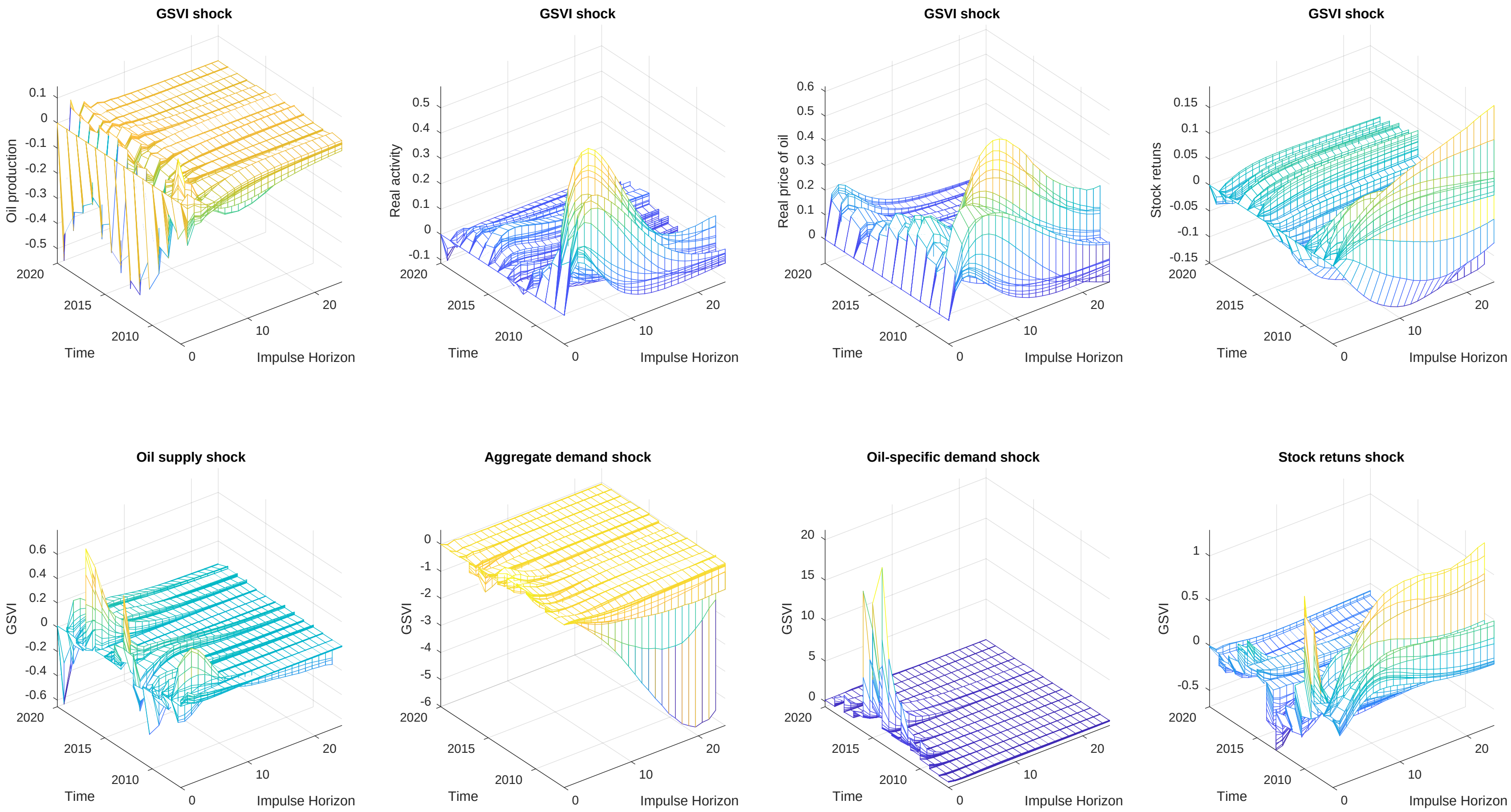


5.2 The role of oil inventories

The impact of oil inventories, through speculative trading, is well documented in the literature (see for instance Hamilton, 2009a,b; Kilian and Murphy, 2014; Baumeister and Hamilton, 2019; Känzig, 2020). However, as noted in Caldara et al. (2019) inventories could in principle quickly move to absorb differences between oil production and oil consumption, in turn affecting the dynamics of the oil market. An extension of our empirical analysis relates to the role of inventories, since our baseline model assumes that oil production is absorbed by consumption in every period. Thus, as an additional robustness check, we expand the five variable model by including a proxy for the oil inventories.

Our aim is to examine whether our findings are affected by the inclusion of oil inventories in the model. Doing so, we repeat the impulse response function analysis regarding shocks to and from the GSVI. These results are reported in Figure 9. In the first row, we plot the responses of the change in oil production, real economic activity, oil prices and stock market returns to a positive shock in information demand. The responses of the variables to a shock in the GSVI are similar to the responses in the main part of the analysis. In addition, the shock leads to an increase in oil demand which yields a positive shift in the price of oil. The increase both in oil demand and the price of oil is greater during the earlier period of the sample, something which we do not observe in the main results. The sign of the response of the stock market changes over time. With few exceptions, stock market returns decline if the shock occurs before 2015 and rise if the shock occurs after 2015. The second row of Figure 9 presents the impulse responses to supply and demand shocks. Similar to the main findings, the sign of the responses and the response of GSVI is affected mostly by the time of the shock.

Figure 9: Impulse responses to shocks from and to the GSVI. The model includes a proxy for oil inventories.



We subsequently consider the effect of oil inventories. In Figure 10 we plot the responses of all variables to a speculative shock. An unanticipated increase in oil inventory demand yields an increase in oil prices which results to a decrease in economic activity. Furthermore, a speculative shock leads to a rise in oil production which lasts only for one period. These findings are in line with the findings of Baumeister and Hamilton (2019). The speculative shock has a negative effect on stock market returns, a result also suggested by Ahmadi et al. (2016). The response of the GSVI is not statistically significant suggesting that a rise in inventory demand does not imply a rise in information demand. Figure 11 presents the impulse response of oil inventory to different shocks. Oil inventories do not respond to an increase in oil production but decline after a positive oil demand shock. This finding contradicts the results from Kilian and Murphy (2014) where the effect of a demand shock on impact increases oil inventories. However, this could be due to the sign restrictions, since in the next periods, the response of oil inventories becomes negative. An increase in prices reduces the oil stocks but the effect is rather weak. On the contrary, a positive shock in stock market has a strong impact on oil inventories. The effect is mostly negative, however for specific time periods, i.e. 2007-2009 and 2013-2014, the shock can lead to an increase in oil inventories. Finally, an unexpected increase in information demand leads to an decrease in oil inventories one month after the shock. The effect is reversed over the next few months. If the shock occurs from 2015 to 2017, the period after the sudden reduction of oil prices, the effect of an information demand shock on oil inventories is reversed.

Figure 10: Impulse responses of the variables to a speculative shock.

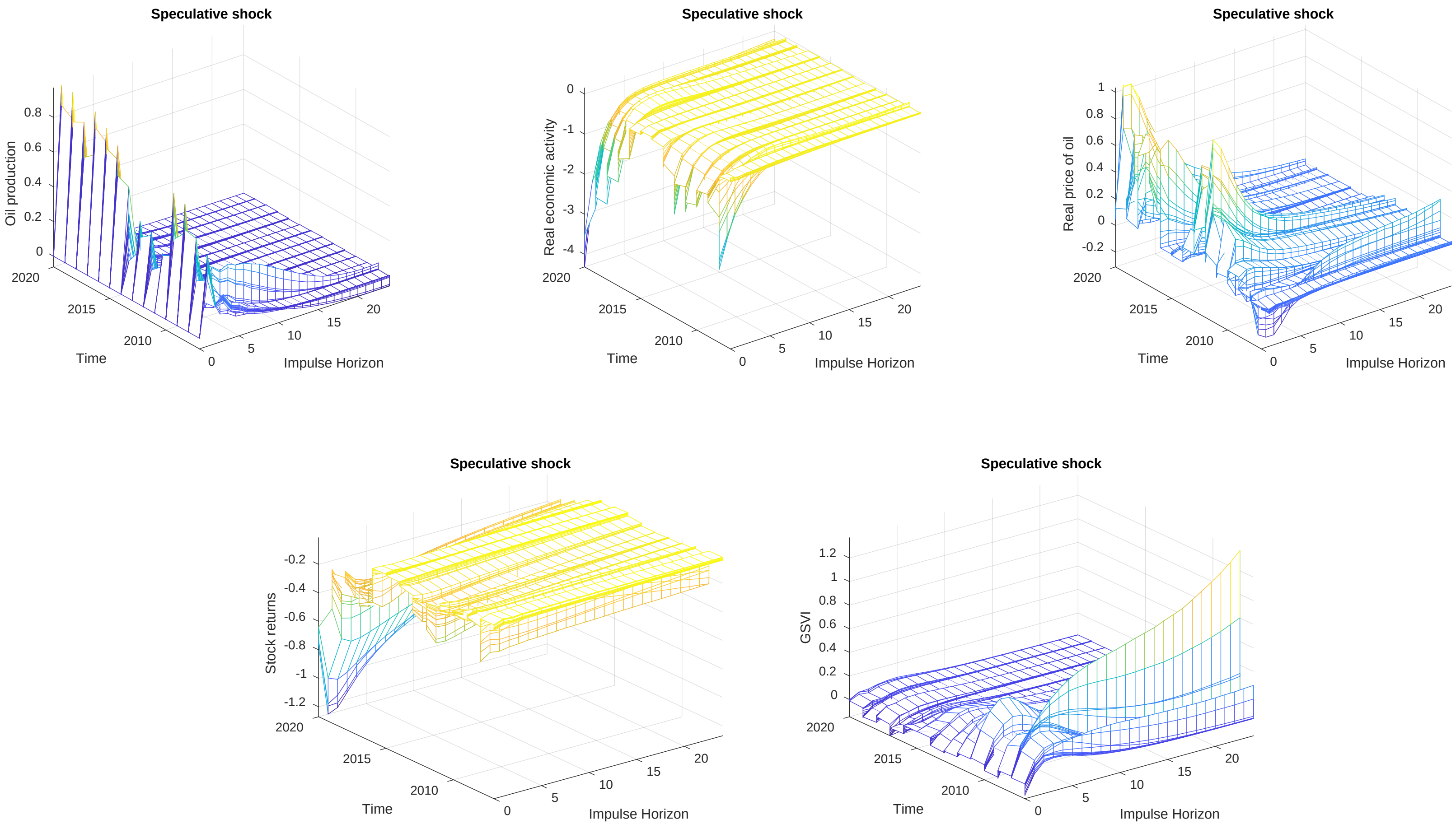
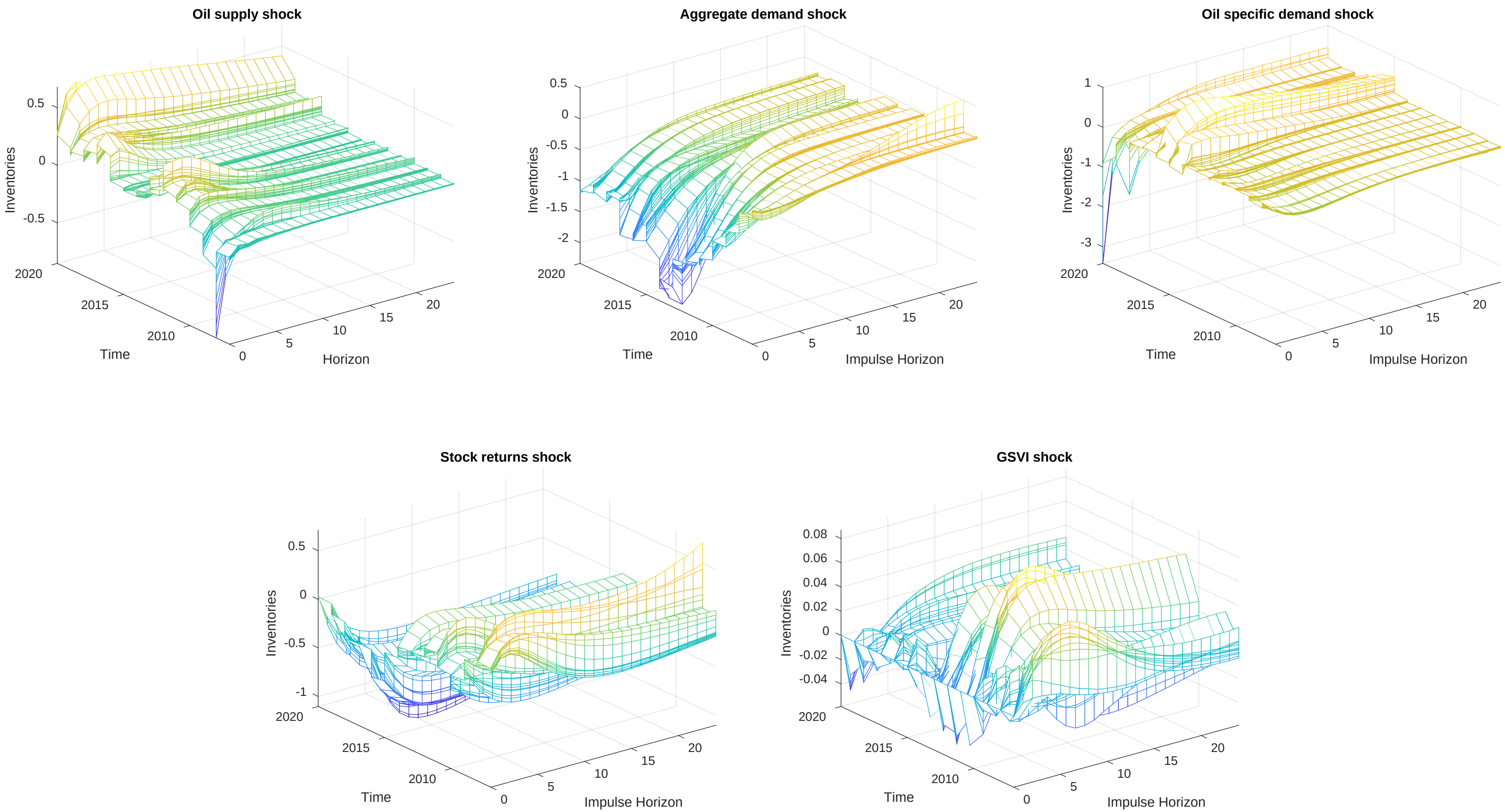


Figure 11: Impulse responses of oil inventories to alternative shocks.



6 Conclusions

The empirical role of retail investors' sentiment shocks as a driver of oil price fluctuations remains debated in the literature, with findings hinging upon the identification assumptions used. This paper examines whether behavioural factors have an impact on oil prices. In order to get a direct measure of investor sentiment, we construct a Google search volume index (GSVI) by employing a dynamic factor analysis on 186 oil related search terms. The sentiment index is added to the model of Kilian and Park (2009) and the impulse response function analysis reveals that a sudden increase in information demand increases the price of oil through an increase in aggregate demand. Furthermore, the sign of the response of stock market returns to the shock depends on the timing of the shock. For example, a shock in the GSVI yields negative stock market returns before 2015 and positive stock market returns after 2015.

Motivated by previous evidence that oil price shocks are transmitted mainly through the demand side, we examine the effect of structural oil demand and supply shocks on GSVI —a barometer of households' and retail investors perception of uncertainty and current and future economic conditions. We show that the aggregate oil supply and demand shocks, oil-specific demand shocks and stock returns shocks have mostly a short-lived and significant impact on GSVI. The impact of aggregate oil supply and demand shocks is found similar in magnitude with the sign of the response to change over time only for the demand side.

Furthermore, economic policy uncertainty (EPU) plays also an important role in influencing financial and economic activities and may attract the attention of investors in the oil market, thus affecting the relationship between investor attention and oil market prices. By adding the EPU index in the model, we find that the effect of the EPU index on the variables is important but short-lasting. However, the EPU responds strongly to shocks in oil prices and stock market returns. In both examined models the results indicate that both the coefficients and transmission mechanisms of the shocks change over time. Finally, our findings are robust to an alternative measure of investor sentiment and after taking into account the role of oil inventories.

References

Ahmadi, M., Manera, M., and Sadeghzadeh, M. (2016). Global oil market and the U.S. stock returns. *Energy*, 114:1277–1287.

Antonakakis, N., Chatziantoniou, I., and Filis, G. (2014). Dynamic spillovers of oil price shocks and economic policy uncertainty. *Energy Economics*, 44:433–447.

Apergis, N. and Miller, S. (2009). Do structural oil-market shocks affect stock prices? *Energy Economics*, 31(4):569–575.

Bai, J. and Ng, S. (2002). Determining the number of factors in approximate factor models. *Econometrica*, 70(1):191–221.

Bai, J. and Ng, S. (2007). Determining the number of primitive shocks in factor models. *Journal of Business & Economic Statistics*, 25(1):52–60.

Baker, S., Bloom, N., and Davis, S. (2016). Measuring economic policy uncertainty. *The Quarterly Journal of Economics*, 131(4):1593–1636.

Barberis, N., Shleifer, A., and Vishny, R. (1998). A model of investor sentiment. *Journal of Financial Economics*, 49(3):307–343. Reprinted in Richard Thaler, ed., Advances in Behavioral Finance Vol. II, Princeton University Press and Russell Sage Foundation, 2005.

Baumeister, C. and Hamilton, J. (2019). Structural interpretation of vector autoregressions with incomplete identification: Revisiting the role of oil supply and demand shocks. *American Economic Review*, 109(5):1873–1910.

Baumeister, C. and 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.

Baumeister, C. and Peersman, G. (2013). Time-varying effects of oil supply shocks on the us economy. *American Economic Journal: Macroeconomics*, 5(4):1–28.

Bird, R. and Yeung, D. (2012). How do investors react under uncertainty? *Pacific-Basin Finance Journal*, 20(2):310–327.

Blanchard, O. J. and Riggi, M. (2013). Why are the 2000s so different from the 1970s? a structural interpretation of changes in the macroeconomic effects of oil prices. *Journal of the European Economic Association*, 11(5):1032–1052.

Caldara, D., Cavallo, M., and Iacoviello, M. (2019). Oil price elasticities and oil price fluctuations. *Journal of Monetary Economics*, 103:1–20.

Carroll, C. D., Fuhrer, J. C., and Wilcox, D. W. (1994). Does consumer sentiment forecast household spending? if so, why? *The American Economic Review*, 84(5):1397–1408.

Carter, C. and Kohn, R. (1994). On Gibbs sampling for state space models. *Biometrika*, 81(3):541–553.

Cogley, T. and Sargent, T. (2002). Evolving post-world war ii u.s. inflation dynamics. In *NBER Macroeconomics Annual 2001, Volume 16*, pages 331–388. National Bureau of Economic Research, Inc.

Cogley, T. and Sargent, T. (2005). Drift and Volatilities: Monetary Policies and Outcomes in the Post WWII U.S. *Review of Economic Dynamics*, 8(2):262–302.

Da, Z., ENGELBERG, J., and GAO, P. (2011). In search of attention. *The Journal of Finance*, 66(5):1461–1499.

Da, Z., Engelberg, J., and Gao, P. (2015). The sum of all FEARS investor sentiment and asset prices. *Review of Financial Studies*, 28(1):1–32.

Degiannakis, S., Filis, G., and Floros, C. (2013). Oil and stock returns: Evidence from european industrial sector indices in a time-varying environment. *Journal of International Financial Markets, Institutions and Money*, 26(10):175–191.

Del Negro, M. and Primiceri, G. (2015). Time-varying structural vector autoregressions and monetary policy: A corrigendum. *The Review of Economic Studies*, 82(4):1342–1345.

Doz, C., Giannone, D., and Reichlin, L. (2011). A two-step estimator for large approximate dynamic factor models based on Kalman filtering. *Journal of Econometrics*, 164(1):188–205.

Doz, C., Giannone, D., and Reichlin, L. (2012). A quasi—maximum likelihood approach for large, approximate dynamic factor models. *The Review of Economics and Statistics*, 94(4):1014–1024.

Elder, J. and Serletis, A. (2010). Oil price uncertainty. *Journal of Money, Credit and Banking*, 42(6):1137–1159.

Frühwirth-Schnatter, S. (1994). Data augmentation and dynamic linear models. *Journal of Time Series Analysis*, 15(2):183–202.

Galay, G. (2019). Are crude oil markets cointegrated? testing the co-movement of weekly crude oil spot prices. *Journal of Commodity Markets*, 16(C).

Guntner, J. H. F. and Linsbauer, K. (2018). The Effects of Oil Supply and Demand Shocks on U.S. Consumer Sentiment. *Journal of Money, Credit and Banking*, 50(7):1617–1644.

Hamilton, J. (2009a). Causes and consequences of the oil shock of 2007-08. *Brookings Papers on Economic Activity*, 40(1):215–283.

Hamilton, J. (2009b). Understanding crude oil prices. *The Energy Journal*, 30(2):179–206.

Hamilton, J. (2013). Historical oil shocks. In Parker, R. and Whaples, R., editors, *The Routledge Handbook of Major Events in Economic History*, pages 239–265. Routledge Taylor and Francis Group, New York.

Hamilton, J. (2021). Measuring global economic activity. *Journal of Applied Econometrics*, 36(3):293–303.

Hong, H. and Stein, J. C. (1999). A unified theory of underreaction, momentum trading, and overreaction in asset markets. *The Journal of Finance*, 54(6):2143–2184.

Jordá, O. (2005). Estimation and inference of impulse responses by local projections. *The American Economic Review*, 95(1):161–182.

Kang, W. and Ratti, R. (2013). Oil shocks, policy uncertainty and stock market return. *Journal of International Financial Markets, Institutions and Money*, 26(3):305–318.

Kang, W., Ratti, R., and Yoon, K. (2015). Time-varying effect of oil market shocks on the stock market. *Journal of Banking & Finance*, 61:150 – 163.

Kilian, L. (2008). The economic effects of energy price shocks. *Journal of Economic Literature*, 46(4):871–909.

Kilian, L. (2009). Not all oil price shocks are alike: Disentangling demand and supply shocks in the crude oil market. *American Economic Review*, 99(3):1053–1069.

Kilian, L. (2019). Measuring global real economic activity: Do recent critiques hold up to scrutiny? *Economics Letters*, 178(3):106–110.

Kilian, L. and 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.

Kilian, L. and Park, C. (2009). The impact of oil price shocks on the U.S. stock market*. *International Economic Review*, 50(4):1267–1287.

Kilian, L. and Zhou, X. (2018). Modeling fluctuations in the global demand for commodities. *Journal of International Money and Finance*, 88:54–78.

Kim, S., Shephard, N., and Chib, S. (1998). Stochastic volatility: Likelihood inference and comparison with arch models. *The Review of Economic Studies*, 65(3):361–393.

Koop, G. and Korobilis, D. (2009). Bayesian multivariate time series methods for empirical macroeconomics. *Foundations and Trends in Econometrics*, 3(4):267–358.

Korobilis, D. (2017). Quantile regression forecasts of inflation under model uncertainty. *International Journal of Forecasting*, 33(1):11–20.

Kuck, K. and Schweikert, K. (2017). A markov regime-switching model of crude oil market integration. *Journal of Commodity Markets*, 6(C):16–31.

Känzig, D. (2020). The macroeconomic effects of oil supply news: evidence from OPEC announcements. *American Economic Review*, 111(4):1092–1125.

Lagerborg, A., Pappa, E., and Ravn, M. O. (2020). Sentimental business cycles. *Forthcoming in Review of Economic Studies*.

Leduc, S. and Sill, K. (2004). A quantitative analysis of oil-price shocks, systematic monetary policy, and economic downturns. *Journal of Monetary Economics*, 51(4):781–808.

Lee, K. and Ni, S. (2002). On the dynamic effects of oil price shocks: a study using industry level data. *Journal of Monetary Economics*, 49(4):823–852.

Lippi, F. and Nobili, A. (2012). Oil and the macroeconomy: a quantitative structural analysis. *Journal of the European Economic Association*, 10(5):1059–1083.

Matsusaka, J. G. and Sbordone, A. M. (1995). Consumer confidence and economic fluctuations. *Economic Inquiry*, 33(2):296–318.

Plagborg-Moller, M. and Wolf, C. (2021a). Local projection inference is simpler and more robust than you think. *Econometrica*, 89(4):1789–1823.

Plagborg-Moller, M. and Wolf, C. (2021b). Local projections and VARs estimate the same impulse responses. *Econometrica*, 89(2):955–980.

Primiceri, G. (2005). Time varying structural vector autoregressions and monetary policy. *Review of Economic Studies*, 72(3):821–852.

Rahman, S. and Serletis, A. (2011). The asymmetric effects of oil price shocks. *Macroeconomic Dynamics*, 15(S3):437–471.

Sargent, T. and Sims, C. (1977). Business cycle modeling without pretending to have too much a priori economic theory. Working Papers 55, Federal Reserve Bank of Minneapolis.

Sockin, M. and Xiong, W. (2015). Informational frictions and commodity markets. *The Journal of Finance*, 70(5):2063–2098.

Stock, J. and Watson, M. (2016). Dynamic factor models, factor-augmented vector autoregressions, and structural vector autoregressions in macroeconomics. In Taylor, J. and Uhlig, H., editors, *Handbook of Macroeconomics*, volume 2, chapter 8, pages 415–525. Elsevier.

Vlastakis, N. and Markellos, R. (2012). Information demand and stock market volatility. *Journal of Banking & Finance*, 36(6):1808–1821.

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

A Appendix

Table A1: List of searched terms used for the construction of the GSVI.

a barrel of oil	current crude oil price(s)	oil barrel price(s)	oil symbol
barrel of crude oil	current crude oil	oil barrel	oil ticker symbol
barrel of crude	current crude	oil chart(s)	oil ticker
barrel of oil price(s)	current oil price(s)	oil close	oil today
barrel of oil	current oil	oil closing price	oil trading price
barrel oil price(s)	current price of a barrel of oil	oil commodities	oil trading
barrel oil	current price of crude oil	oil commodity	price barrel of oil
barrel price(s)	current price of crude	oil contract	price barrel oil
bloomberg crude	current price of oil	oil cost per barrel	price crude oil
bloomberg energy prices	daily oil price(s)	oil cost	price crude
bloomberg energy	energy fund(s)	oil crude	price for barrel of oil
bloomberg oil	energy index	oil current price	price of a barrel of oil
buy oil futures	energy mutual funds	oil demand	price of barrel of oil
closing oil price(s)	energy prices	oil future	price of barrel
closing price of oil	energy stocks	oil futures price(s)	price of crude oil per barrel
cost of a barrel of oil	future oil	oil futures	price of crude oil
cost of crude oil	futures price	oil graph	price of crude
cost of oil	gas barrel	oil index	price of oil barrel
crude future(s)	gas index	oil market(s)	price of oil
crude oil barrel price	gas per barrel	oil options	price oil barrel
crude oil barrel	gas price per barrel	oil per barrel	price oil
crude oil chart(s)	global oil	oil price barrel	price per barrel of oil
crude oil close	historic oil prices	oil price chart(s)	price per barrel today
crude oil future(s)	historical oil price(s)	oil price graph	price per barrel
crude oil index	history of oil prices	oil price history	real time oil
crude oil market	investing in oil	oil price index	spot crude
crude oil per barrel	latest oil price(s)	oil price per barrel	spot oil price(s)
crude oil price(s)	light crude oil	oil price quote	spot oil
crude oil pricing	light crude	oil price ticker	stock market oil
crude oil quote	light sweet crude oil	oil price today	sweet crude oil
crude oil spot	light sweet crude price	oil price trend	sweet crude price(s)
crude oil stock symbol	light sweet crude	oil price(s)	sweet crude
crude oil stock	live oil price	oil prices graph	symbol for crude oil
crude oil symbol	nymex crude future	oil pricing	the price of oil
crude oil ticker symbol	nymex crude oil	oil producers	time oil
crude oil ticker	nymex crude	oil quote(s)	today oil price
crude oil today	nymex oil	oil spot price	trade oil
crude oil	oil a barrel	oil spot	trading oil
crude price(s)	oil and gas prices	oil stock symbol	world oil prices
crude stock	oil barrel cost	oil stocks	
crude	oil barrel price today	oil supply	

Notes: All SVI data are downloaded from Google Trends.

Table A2: Summary statistics of the variables used in the analysis.

Variable / Statistic	Mean	Maximum	Minimum	St. dev.	Skewness	Kurtosis	Jarque-Bera stat.	ADF stat.
Oil production	7356	12859	3973	2425	0.721	2.265	21.36***	-2.283
REA	13.02	191.0	-160.0	78.47	0.360	2.289	8.355**	-3.948**
Oil price	67.04	127.7	16.74	24.38	0.395	2.105	11.65***	-2.923
Stock market	1712	3230	735.0	624.7	0.730	2.375	20.76***	-1.893
GSVI	-0.002	0.463	-3.164	0.548	-2.846	12.38	1066.51***	-2.207
EPU	119.6	283.1	57.20	40.42	0.955	4.250	42.53***	-1.714
ICS	83.98	103.8	55.30	12.18	-0.526	2.224	13.96***	-2.205

Notes: i) ***, ** and * denote rejection of the null hypothesis at the 1%, 5% and 10% significance level. ii) In the implementation of the ADF test we assume only constant in the test equation and for the selection of the lag-length we use the SIC.