DALLASFED Working Paper

OPEC in the News

Michael Plante

Federal Reserve Bank of Dallas
Research Department
Working Paper 1802
https://doi.org/10.24149/wp1802

OPEC in the News *

Michael Plante

Current version: February 2018

First version: April 2015

ABSTRACT

This paper introduces a newspaper article count index related to OPEC that rises in response to important OPEC meetings and events connected with OPEC production levels. I use this index to measure how interest in OPEC varies over time and investigate how oil price volatility behaves when the index unexpectedly changes. I find that unexpected increases in the newspaper index are strongly associated with higher levels of oil price volatility, both realized and implied. In some cases, interest levels and price volatility appear to be driven by the OPEC event itself, such as the Iraq invasion of Kuwait. In other cases, such as the oil price collapses in late 2008 and late 2014, price volatility and interest levels in an OPEC event appear to be responding endogenously to developments in the oil market or broader economy. The newspaper index is highly correlated with Google search volume data on OPEC, an alternative measure of the amount of attention paid to OPEC events.

Keywords: Oil price volatility; OPEC; Attention; Google search activity
JEL Classifications: Q43; C32; D80; G12

*Plante, Research Department, Federal Reserve Bank of Dallas, 2200 N Pearl St, Dallas, TX 75201 (michael.plante@dal.frb.org). This paper previously circulated under the title "OPEC in the News: The Effect of OPEC on Oil Price Uncertainty." I am grateful to Nathan Balke, Alex Chudik, Everett Grant, Soojin Jo, Alex Richter and participants of the Dallas Fed - Norges Bank joint workshop on applied econometrics for helpful comments on this version. I also thank Justin Lee, Grant Strickler and the Dallas Fed library staff for excellent research assistance. All errors are my own. The views expressed in this paper are those of the author and do not necessarily reflect the views of the Federal Reserve Bank of Dallas or the Federal Reserve System.

1 INTRODUCTION

A growing literature has made use of news coverage and internet search volume data to analyze economic phenomena. Barber and Odean (2008) [9], for example, use news coverage data to help explain the decisions of individual investors. Da, Engleberg and Gao (2011) [11] use Google Search Volume Index (SVI) data to help explain stock price movements. Andrei and Hassler (2014) [4] use Google search volumes to study the relationship between investors' attention levels and stock market volatility. Baker, Bloom and Davis (2016) [7], hereafter BBD, introduced a monthly index based on newspaper counts to track the level of economic policy uncertainty. Afkhami, Cormack and Ghoddusi (2017) [2] show that weekly Google SVI data on numerous keywords related to energy have predictive power for energy price volatility. Baker and Fradkin (2017) [8] use Google search data to measure job search activity.

There is also an extended literature that has explored how OPEC and OPEC-related events affect oil price volatility. Horan, Peterson and Mahar (2004) [19], for example, use an event study approach and find that market expectations of crude oil price volatility, given by an option-implied volatility, often rises in the days before an OPEC meeting but falls thereafter. Robe and Wallen (2016) [27] find that OPEC spare capacity has explanatory power for the implied volatility of oil. A number of other works have also documented a connection between OPEC and different measures of oil price volatility.

In this paper, I make use of news coverage data to consider the connections between OPEC and oil price volatility, the first work to do so to the best of my knowledge. A contribution of this work is to introduce a monthly newspaper article count variable that is used to measure how interest in OPEC and OPEC events varies over time. This index, hereafter referred to as the OPEC newspaper index, is similar in spirit to the economic policy uncertainty index of BBD but is based on the number of newspaper articles with references to OPEC. My hypothesis is that the number of articles written about OPEC should vary from month to month depending upon the amount of attention being paid to the group. During periods of intense scrutiny about OPEC, for example, there should be more articles written about the group and its members. This provides information not only on what events are considered important but also on how attention levels vary over time. I show that this index does indeed vary over time, and rises in response to a variety of important events related to OPEC, such the oil price collapse in 1986, the Iraq invasion of Kuwait and select OPEC meetings.

I then use the index to explore how oil price volatility behaves around periods of unusual interest in the cartel. To do so, I incorporate the newspaper index into a variety of reduced-form, monthly vector autoregression (VAR) models and consider what happens when there is an unexpected change in the OPEC newspaper index.[1] To avoid making strong, and potentially incorrect, identifying assumptions about causality I use generalized impulse response functions to do the analysis. In this case, the shocks to the newspaper index are the residuals from its regression equation, i.e. the unforecastable changes in the index.

I begin with simple bi-variate VAR models that contain the newspaper index and either a measure of realized volatility or implied volatility for West Texas Intermediate (WTI) crude oil. I find that there is a very strong statistical connection between unexpected increases in the newspaper

[1] Monthly frequencies allow the newspaper index to be added into models that contain oil supply and demand data, variables known to be important for modeling the oil market but unavailable at daily or weekly frequencies. See Kilian (2009) [22] and related papers for recent examples.

index and higher levels of oil price volatility, both realized and implied. Several robustness checks then are undertaken to test the strength of these findings. In one check, the empirical models are expanded to include data on global oil production, global oil demand, and oil prices, to control for the possibility that movements in the newspaper index might be predictable given lagged values of these variables. Adding these variables has very little impact on the results. The results are also robust to using Brent crude price and volatility data in place of West Texas Intermediate crude, and to using an alternative formulation of the newspaper index that takes into account changes in the volume of total newspaper articles that have occurred over time.

The OPEC newspaper index is constructed using a simple search method based on newspaper counts. This approach may be subject to concerns regarding its accuracy and reliability in measuring interest levels in OPEC. To alleviate these concerns, I compare the newspaper index to Google SVI data that tracks search frequency on the topic of OPEC, available since 2004. I find the OPEC newspaper index is highly correlated with SVI data, with a correlation coefficient of 0.78.[2] The empirical findings in the VAR are also robust to replacing the newspaper index with Google SVI data. While highly correlated with the Google data, the OPEC newspaper index has several advantages over it. One is that it extends back to 1986, whereas Google data is only available starting in 2004. The index is also easily obtained over a long period of time in different frequencies, including weekly, monthly, quarterly or annually.

I also investigate what particular events correspond with unexpectedly high interest levels in OPEC. To do this, I use a narrative-type approach and examine the residuals from the empirical model to see whether they line up with an OPEC-related event, focusing on the largest shocks. The largest unexpected increase occurred in August 1990 when Iraq invaded Kuwait. The second largest was in October 2014, when a rapid decline in oil prices fueled unusual interest in OPEC's November 2014 meeting. Other notable, unpredicted increases occurred in August 1986 during an OPEC meeting in Geneva that led to a surprise production agreement; OPEC meetings in 1992 and 1993 during the return of Kuwaiti production; OPEC meetings in late 1997 and early 1998 during the Asian financial crisis; the post 9-11 meeting in November 2001; and several other meetings in the mid-2000s.

The models and methodology used here do not make any direct assumptions or statements about the causality between the newspaper index and volatility levels. It could be the case that some OPEC events are unusually important, and that this drives attention and volatility levels. Or it could be that unexpectedly high levels of oil price volatility are driven by events in the oil market or broader economy, and these events in turn make OPEC events, particularly meetings, more important and more widely watched than is usually the case. To the extent one can make such statements, the residual analysis suggests that both explanations play a role. For example, the Iraq invasion of Kuwait was an exogenous event that was important in its own right. Likewise, although not exogenous in the sense of a war, the unexpected increase in the index in August 1986 was related to Saudi Arabia's decision to dramatically increase production levels from historic lows in the summer of 1985. On the other extreme, unusual interest in OPEC in October 2014 was clearly related to developments in the oil market. Likewise, the OPEC meeting in October 2008 garnered unusual interest due to recent developments in the economy, i.e. intense concern about demand levels due to the financial crisis and how OPEC might respond to those developments.

[2] A third alternative, Bloomberg News Trends data, is available from 2007 onwards. The newspaper index is also highly correlated with this data.

These findings have some potential implications for thinking about the connections between oil price volatility and the economy, as they suggest that while certain periods of high volatility seem to be due to exogenous OPEC events, other periods appear to be endogenous variations in volatility linked, in part, to events in the broader economy.

This work has connections with the literature that makes use of news and internet search based data to analyze economic phenomena. In addition to previously cited papers, important works in this area include Tetlock (2007) [29], Engelberg and Parsons (2011) [14] and Yuan (2015) [30]. This paper differs from the previous literature in that it uses news and internet search data to measure interest in OPEC events. To the best of my knowledge, this approach has not been used for thinking about how attention levels towards OPEC have varied over time.

This paper also has connections with the literature that focuses on OPEC and oil price volatility. Some important examples include Guo and Kliesen (2005) [18], Demirer and Kutan (2010) [12], Schmidbauer and Rosch (2012) [28], Amelier and Darne (2014) [3] and Mensi, Hammoudeh and Yoon (2014) [24]. My work differs from these in my approach to the question. Many previous works have used event study analysis of various types. Here, I consider the behavior of oil price volatility around those events that attract unexpectedly high levels of interest, as given by the newspaper index or Google SVI data. Such an approach has not been used before in the literature.

Finally, there is also a literature that discusses the connections between oil price uncertainty and economic activity. Bernanke (1983) [10], for example, showed increased uncertainty about oil prices could negatively affect investment spending when there is an irreversible investment decision. Numerous empirical works have also found evidence of a negative relationship between various measures of oil price uncertainty and macroeconomic activity. Notable examples include Ferderer (1996) [15], Guo and Kliesen (2005) [18], Elder and Serletis (2010) [13], and Jo (2013) [20], amongst many others. Kellogg (2014) [21] shows that oil price uncertainty affects the investment decisions of oil and gas companies.

The rest of the paper is organized as follows. In the next section I introduce the newspaper index and other data used in the analysis. In Section 3 I discuss the empirical methods while in Section 4 I present the results. Section 5 concludes.

2 DATA

2.1 OPEC NEWSPAPER INDEX

I am interested in measuring how attention paid to OPEC and OPEC events varies over time. To do so, I construct an article count variable similar in spirit to the economic policy uncertainty index in BBD (2016) [7]. The article count variable, hereafter referred to as the OPEC newspaper index, is constructed using monthly article counts from several major newspapers. The use of monthly data allows the index to be incorporated into empirical models of the oil market, which are often monthly in frequency due to limitations on supply and demand data. It is possible to generate the index at other frequencies, as well, such as weekly or quarterly. The time series for the monthly index is available online.[3]

The source for the raw counts is the Dow Jones Factiva database. The search is based on the keyword "OPEC" and generates a monthly count for four major newspapers: the Financial Times, the Houston Chronicle, the New York Times and the Wall Street Journal. The appendix provides the exact search criteria. These four papers were chosen given the fact that they are widely read, have many articles devoted to the economy, and devote space to writing about the oil market. As a

[3]The data is available at sites.google.com/site/michaelplanteecon/research.

result, major events related to OPEC that capture the interest of oil market participants are likely to be written about in these papers. The index begins in 1986, the first year that Factiva provides full data for the Houston Chronicle. It is possible to go back to the early 1980s using the other three papers. The series could be extended even further given access to more detailed newspaper records than those available with Factiva. The search finds a total of 22,183 articles over the full sample period. The average number each month is 59.6 articles per month, while the median count is 46.

A first step to creating the index is to aggregate across newspapers. One issue is that the counts differ across newspapers, so simply summing across the raw article counts could overweight the influence of papers with systematically larger article counts. To deal with this, I follow BBD and first standardize the individual newspaper time-series to unit standard deviation. More specifically, for each newspaper i I calculate the sample standard deviation, σ_i, for the full sample from January 1986 to December 2016. I then divide the time series for newspaper i by σ_i. This provides four time series that provide easily interpretable information on when there are unusually large number of articles being written about OPEC in each specific paper. Table 1 shows the correlations between the four measures. Most are highly correlated with each other, although the correlations between the FT and the other papers are somewhat weaker.

Financial Times	1			
Houston Chronicle	0.57	1		
New York Times	0.64	0.86	1	
Wall Street Journal	0.72	0.84	0.85	1

Table 1: Correlations between the newspaper article counts. The sample runs from Jan. 1986 to Dec. 2016.

To construct the OPEC newspaper index, I then perform the following calculations. For each month, I first average across the four standardized time-series. This new series is denoted as Z_t. Next I calculate the mean of this series, denoted by M. Finally, I multiply the Z_t by $100/M$, for all months. This produces a normalized series with an average of 100 over the sample period considered.

Figure 1 plots the OPEC newspaper index. The largest spike coincides with the Iraq invasion of Kuwait in August 1990, which itself was preceded by a very large jump in July 1990. There are a number of spikes in 1986 following Saudi Arabia's decision to rapidly increase production levels from historic lows in the summer of 1985. Other notable spikes in 1987 and 1988 line up with OPEC meetings. Large spikes in March 2000 and September 2000 line up with OPEC meetings, and November 2001 lines up with an emergency OPEC meeting had after 9/11. Increases in late May 2004 picks up an OPEC meeting and militant attacks that occurred in Saudi Arabia. Later in the sample, the index picks up OPEC meetings in October 2006, November 2007, emergency meetings in October and December 2008, plus meetings at the end of 2014 and 2016. While the index typically rises contemporaneously with an event, it can display some forward looking behavior and rise in advance of an event, such as the increase in October 2014. Overall, there is strong evidence that the OPEC newspaper index is closely linked to events related to OPEC, typically meetings but occasionally other types of events as well.

Many increases in the index line up with the bi-annual OPEC conferences. The dates for these meetings are publicly known well in advance and it seems quite plausible that the number

of articles written about OPEC would naturally increase during those months, relative to other months. This would create a type of irregular, but predictable, seasonal pattern in the index. I have tested for this possibility by regressing the index on a constant and a dummy variable that is set to 1 during a month with one of the bi-annual conferences. The results overwhelmingly confirm the hypothesis that there is an increase in counts during those months, with the coefficient on the dummy variable having a p-value near zero. The empirical analysis, details of which are introduced in the next section, controls for this predictable increase in the index.

2.1.1 ALTERNATIVE INDEX In the baseline newspaper index, it is the raw number of articles written about OPEC that is considered informative about the level of interest in the group. One issue with this approach is that print edition newspapers, in general, have been shrinking in size over time. As a result, the number of total articles being written about all subjects has been declining. To take this into account, I construct an alternative index that takes the total number of articles written about OPEC across the four papers each month and divides it by the total number of articles produced by the four papers that month.[4] A priori, it is not clear whether the baseline or alternative approach might be preferred. While the alternative index does control for trends in the overall number of articles that appear in newspapers, it introduces noise because the total number of articles can vary from month to month for any number of reasons. In any case, I find that the alternative index is closely related to the baseline index: the correlation coefficient between the two is 0.97.

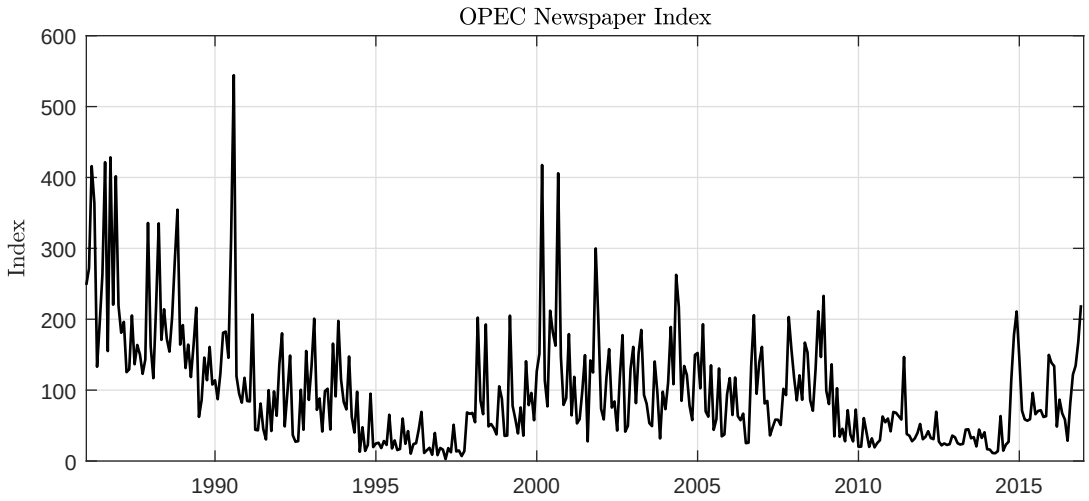

Figure 1: Time series of the OPEC newspaper index. The index values average 100 over the full sample.

[4]The appendix describes the search criteria used to generate the estimate of the overall volume of articles.

2.2 GOOGLE SEARCH VOLUME INDEX One alternative to using news coverage data to gauge interest in a topic is data based on internet search volumes. Google Search Volume Index (SVI) data, available from Google Trends, has been frequently used for this purpose in a number of other research papers.[5] This data allows one to measure search intensity for a particular keyword or topic, over a specified period of time and within a geographic area.

There is a specific series that tracks search volumes for the topic of "OPEC" that incorporates searches based on several different keywords. Data is available from 2004 to the present and comes in the form of an index that ranges from 0 to 100. The index is normalized such that 100 represents the period of maximum search intensity over the sample period being considered.

I extract monthly SVI data from 2004 to 2016 based on worldwide searches related to the topic OPEC. Figure 2 plots the time series. The SVI reaches a maximum on May 2004, which lines up with the militant attacks in Saudi Arabia. Several other notable periods with high values include November 2007 (OPEC meeting), May 2008 (militant attacks in Nigeria), September - December 2008 (financial crisis and OPEC meetings), November 2014 (OPEC meeting) and late 2016 (OPEC meeting). In general, the SVI and the newspaper index contemporaneously respond to many of the same events, but not all. The correlation coefficient between the two is 0.78.

The monthly SVI data has a visible seasonal pattern, with a decline typically occurring in the summer in July. The reason for this is unclear but given that it exists I seasonally adjust the SVI data before using it in the empirical analysis.[6] There is also evidence that the SVI data rises predictably during the regularly scheduled bi-annual OPEC meetings, which I control for in the empirical analysis.

The newspaper index and SVI data each have their own pros and cons in terms of measuring attention levels towards OPEC events. One useful feature of the SVI data is that it is, as stated by Da et al. (2011) [11], a *revealed* attention measure. That is, people made an explicit attempt to search for OPEC on Google. A major advantage of the newspaper index is that it can be extended much further back in time than SVI data. The current index starts in 1986, the first full year Factiva provides data on the four newspapers used in the analysis.

2.3 OIL MARKET DATA The empirical analysis also makes use of data on world oil production, world industrial production (as a proxy for oil demand), crude oil prices, a measure of realized volatility for crude oil prices and an implied volatility series for crude oil. I now describe this data in more detail.

The oil production data is the Energy Information Administration's world crude oil production including lease condensate series. I use the world industrial production series from the Dallas Fed's Database on Global Indicators as a proxy for oil demand. This series covers industrial production from a large number of countries.[7] My oil price variable is the inflation-adjusted spot price of West Texas Intermediate crude oil. I take a monthly average of daily nominal spot prices available from Bloomberg and deflate the monthly average using the U.S. CPI excluding food and energy prices. I transform the supply, industrial production and real oil price data by taking logs, differencing, and multiplying by 100. The sample for these data runs from January 1986 to December 2016, providing a total of 372 observations.

[5] See for example Da, Engleberg and Gao (2011) [11], Afkhami, Cormack and Ghoddusi (2017) [2] and Baker and Fradkin (2017) [8].

[6] The empirical results are not qualitatively affected by seasonally adjusting the data.

[7] See Grossman, Mack and Martinez-Garcia (2014) [17] for more details.

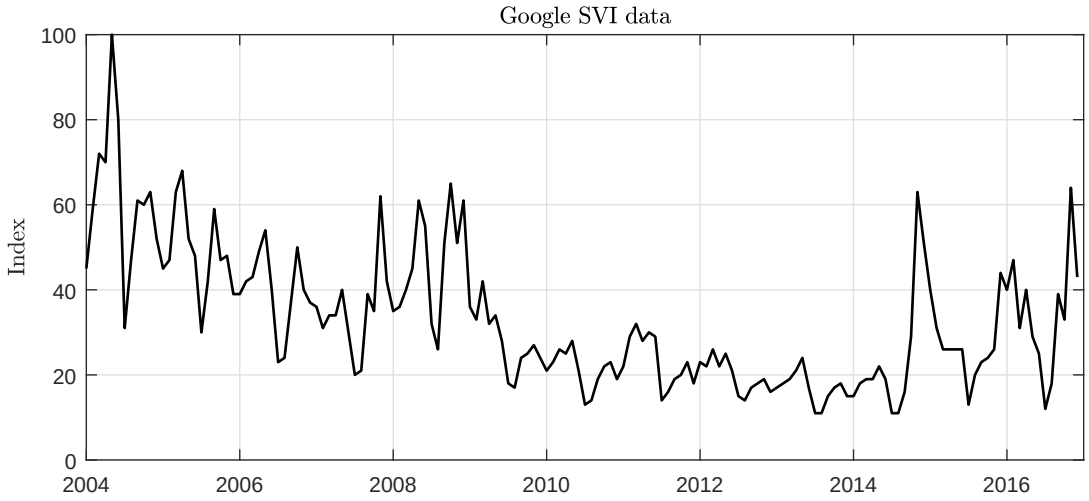

Figure 2: Time series of Google Search Volume Index data on OPEC.

I use daily spot prices for West Texas Intermediate crude oil to derive a realized volatility series. The price on any given day in month t is denoted as $P_{d,t}$. The daily return on day d in month t is calculated as

$$r_{d,t} = \ln\left(\frac{P_{d,t}}{P_{d-1,t}}\right). \tag{1}$$

For each month, I calculate an average return which I denote as $\bar{r}_t$. The monthly volatility, denoted as HV_t for month t is calculated as

$$HV_t = 100 \times \sqrt{\frac{260}{nday_t - 1}\sum_{d=1}^{nday_t}\left(r_{d,t} - \bar{r}_t\right)^2}, \tag{2}$$

where $nday_t$ is the number of trading days in month t.

The implied volatility series is from options on the one-month ahead futures contract for West Texas Intermediate (WTI) crude oil. The series measures the market's expectations of near-term volatility in the nominal price of WTI. Data from Bloomberg provides a daily series of implied volatilities starting in July 1993. A monthly series is generated by averaging across the daily observations in each month. The series runs July 1993 to December 2016 for a total of 282 observations.

3 METHODS

3.1 MODELS The empirical models are monthly vector autoregression (VAR) models given by

$$x_t = c + \Psi D_t + \sum_{i=1}^{p} \Phi_i x_{t-i} + \epsilon_t, \tag{3}$$

where p is the lag length. The baseline models are bi-variate and include the newspaper index and a measure of oil price volatility. The vector x_t is a 2×1 vector of jointly determined variables, c is a 2×1 vector of constants, Φ_i are 2×2 matrices parameters, and ϵ_t is a 2×1 vector of residuals. The D_t term is a deterministic dummy variable set to 1 during months with one of the bi-annual OPEC conference meetings, while Ψ is a 2×1 vector. I denote Σ as the variance-covariance matrix of the residuals. The lag length p is determined using the Akaike information criterion (AIC). In the robustness check section, a set of larger VAR models are also considered that include oil production, demand and prices as variables in the model.

3.2 IMPULSE RESPONSE FUNCTIONS To examine the effects of an unexpected increase in the newspaper index on oil price volatility I employ generalized impulse response functions, developed by Pesaran and Shin (1998) [26]. This section provides a brief overview of the method.

The generalized impulse response function for variable $x_\ell, \ell = 1, ..., k$, at time t in response to a shock to variable j is given by

$$GI_\ell\left(n, \delta_j, \Omega_{t-1}\right) = E(x_{\ell,t+n}|\epsilon_{jt} = \delta_j, \Omega_{t-1}) - E(x_{\ell,t+n}|\Omega_{t-1}), \tag{4}$$

where $n = 0, 1, 2, ...,$ is the forecast horizon, $\Omega_{t-1} = \{x_{t-1}, x_{t-2}, ...\}$ is the information set of all lagged observables and δ_j is the size of the shock.

In the context of the model in (3), my question is analogous to asking what happens if the residual in the newspaper index equation is non-zero at time t. The exercise considers a positive, one-standard deviation size shock. Equation (4) then provides the answer about how expectations regarding all of the variables in the model should be adjusted due to this shock, relative to what would happen if there were no shock at all.

Generalized impulse response functions are used as I do not wish to take a strong stance on identifying what underlying structural shocks might be behind a particular unforecastable change in the newspaper index. While identifying causal connections is of significant interest, the strong assumptions necessary to do so may give mis-leading answers about causality if those assumptions are incorrect. A more detailed discussion on causality is found in section 4.3.2.

4 RESULTS

4.1 BI-VARIATE VARS I begin the analysis with a pair of reduced-form, bi-variate VAR models that include just the log of the newspaper index and a measure of oil price volatility. The first model uses the realized volatility series for WTI prices given by equation (2). The sample for this model runs from January 1986 to December 2016, a total of 372 observations. The lag length is set to 4 lags, based on an AIC test done with 12 lags as the maximum lag length considered. The second model uses the WTI implied volatility series in place of realized volatility. My data on implied

volatility begins in July 1993, providing 282 observations. I continue to use 4 lags for this model, as well.[8]

I consider a positive, one-standard deviation shock to the newspaper index. Figure 3 plots the responses of the variables. The top panel is for the model with realized volatility, the bottom for the model with implied volatility. For each variable, the solid line is the point estimate for the response, the dotted and dashed lines are one and two standard-error bands, respectively.

For both models, I find that shocks to the newspaper index lead to a persistent increase in the index itself. The initial effect subsides rapidly after one month but remains positive for many months thereafter. The path of the index is basically indistinguishable between models. There is very strong statistical evidence that this shock is associated with higher oil price volatility, whether it be measured by realized volatility or implied volatility. The initial increase is one to two percentage points and the response is statistically significant at the 95 percent level. The increases in volatility are persistent and generally statistically significant over a 12 month horizon.

The impulse responses pick up a strong statistical connection between unexpected changes in the newspaper index and changes in oil price volatility. One explanation for this result is that unexpected increases in the newspaper index occur during a month when there is greater speculation or uncertainty about OPEC production levels. Given that OPEC is a large player in the market and that some members of OPEC are very large producers in their own right, questions about their production levels could certainly generate price volatility. At a deeper level, such unusual speculation could be driven by several different types of events. These could include unexpected supply disruptions in an OPEC country, disagreement between OPEC members about current or future production levels, or events in the oil market or broader economy that raise questions about how OPEC will respond. I discuss these issues in greater detail in Section 4.3 when I look into the events the model picks up as shocks.

4.2 ROBUSTNESS CHECKS I considered several checks to document how robust the previous results are to modifications in the model and data. Overall, I find that the connection between shocks to the newspaper index and oil price volatility are quite robust. Figure 5 plots the generalized impulse response functions for oil price volatility to the newspaper index shock for the different cases considered. The left and right panels plot realized volatility and implied volatility, respectively.

4.2.1 EXPANDED VAR MODEL The bi-variate VAR models provide strong statistical evidence that unexpected changes in the OPEC newspaper index are associated with changes in oil price volatility. However, there are many other variables that have been shown to be important when modeling the oil market. Thus, as a robustness check I now expand the VAR models to include supply, demand and price variables, as in Kilian (2009) [22]. Adding these variables also allows the model to take into account potentially predictable movements in the index due to changes in lags of those three variables. For my models, I specifically use the log difference of world oil production, the log difference of world industrial production, and the log difference of real West Texas Intermediate spot prices. These variables are multiplied by 100. As before, I consider one VAR model that includes WTI realized volatility and another that uses WTI implied volatility.

Figure 4 compares the responses of oil price volatility to a shock to the newspaper index for the small and large VAR models. The left panel shows realized volatility, the right implied volatil-

[8]The AIC test generally picks a lag length of 3 or 4 lags for most of the models considered in this paper.

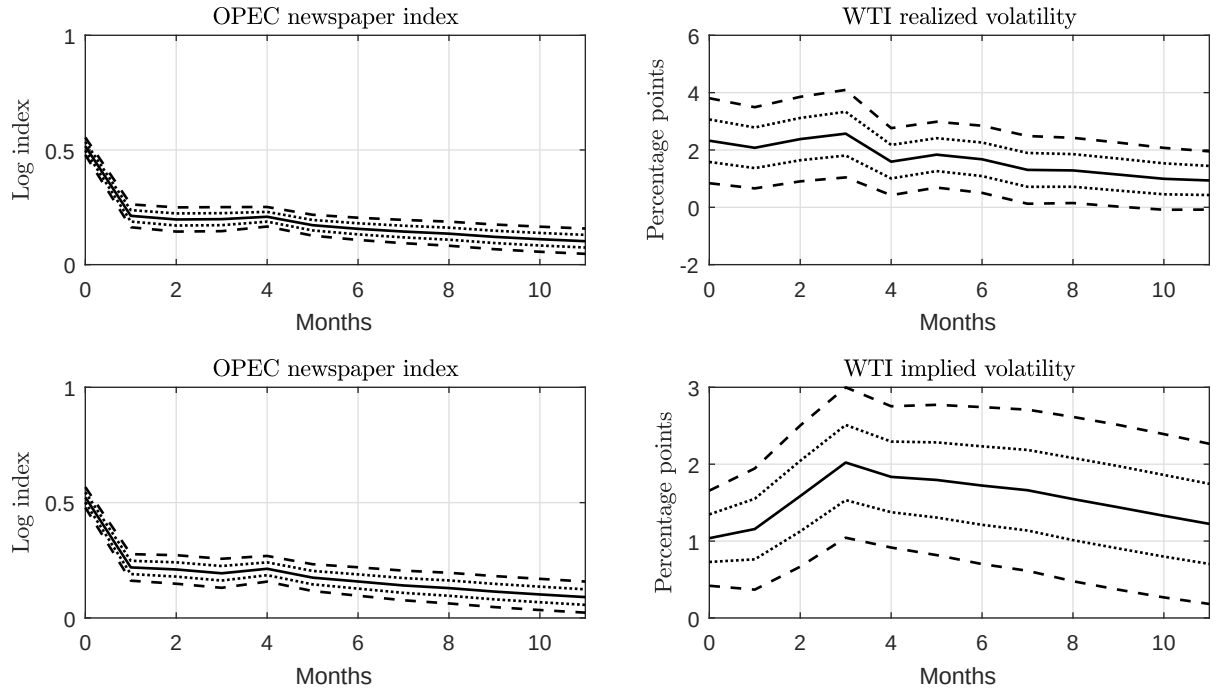

Figure 3: Generalized impulse response functions for a shock to the OPEC newspaper index. Dashed lines are 2 s.e. bands, dotted lines 1 s.e. bands. The top set of figures is for the model with realized volatility. The bottom set is for the model with implied volatility.

ity. Although the model now includes a number of other variables, I find that the responses of the volatility variables are quite similar between the bi-variate case and the larger VAR model. An unexpected increase in the newspaper index continues to be associated with increased price volatility. Although not shown here, I find that the impulse response for the newspaper index itself is very similar across all four models.

4.2.2 BRENT PRICES The second robustness check uses data based on Brent crude as a substitute for WTI price and volatility data in the larger VAR model. For the model that uses realized volatility the sample size remains the same. Data on implied volatility for Brent crude is available starting in 1997, instead of 1993. The dotted lines in Figure 5 show the responses. I find that using Brent crude data does not significantly alter the results.

4.2.3 GOOGLE SVI Another variant replaces the newspaper index with the Google SVI variable. These models are estimated from 2004 to 2016, due to the data limitations on the SVI. These responses are the dashed lines in Figure 5. Overall, I find that unexpected increases in the SVI are associated with higher levels of volatility compared with the models that use the newspaper index.

4.2.4 ALTERNATIVE INDEX The circles in Figure 5 plots the responses when the alternative formulation of the newspaper index is used where the OPEC count is divided by the total volume of articles each month. The circles almost perfectly track the responses of the large VAR that includes the baseline index.

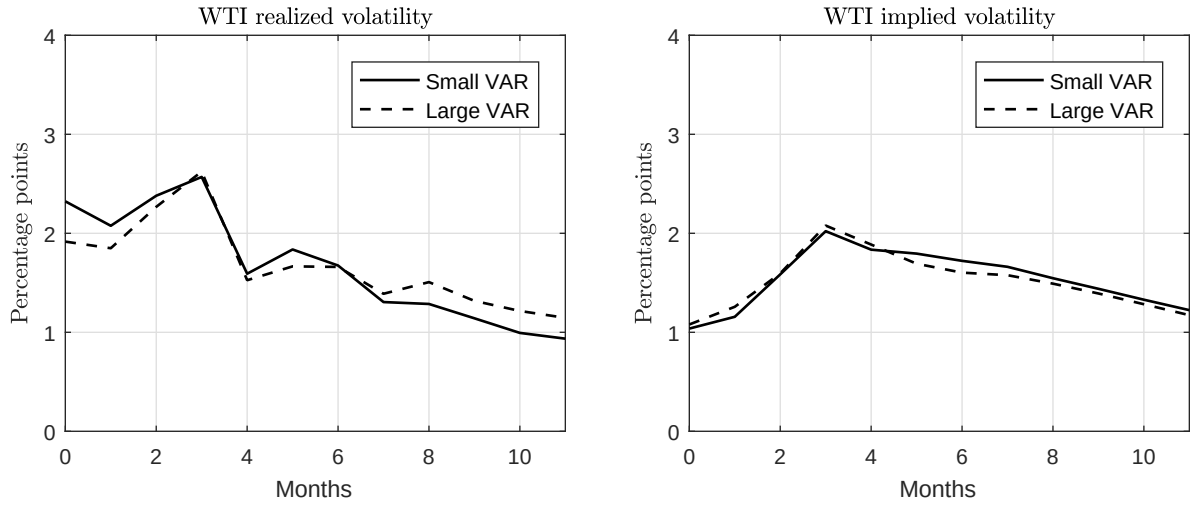

Figure 4: Generalized impulse response functions for a shock to the OPEC newspaper index. The left panel is for VAR models with realized volatility. The right figure for models that include implied volatility.

4.2.5 ALTERNATIVE EXPERIMENT Finally, I also consider an alternative experiment where I ask what happens if there is an unexpected increase in the newspaper index now but no contemporaneous response in any of the other variables. I implement this using a Cholesky decomposition of the VAR model with the newspaper index ordered last. The dashed-dotted line in Figure 5 shows the response of volatility under these assumptions. An unexpected increase in the index foreshadows higher levels of both realized and implied volatility in the near future. The response further out in time is very similar to the responses from the other cases considered.

4.3 ANALYSIS OF RESIDUALS For generalized impulse response functions, the shocks to the newspaper index are the residuals from its equation, i.e. the *unforecastable* changes in the index at time t. Given this, as part of the analysis I investigated the time series of the residuals to see if they corresponded with particular OPEC events. For this analysis, I used the five-variable VAR model with realized volatility. This model allows me to conduct the analysis back to 1986, and the use of the five-variable model ensures that I have eliminated movements in the newspaper index that might be predictable given lagged values of supply, demand and prices. This exercise was also done using Google SVI data but those results are relegated to the appendix, as they were similar in nature to the results for the newspaper index.

4.3.1 NEWSPAPER INDEX Table 2 lists the ten largest shocks from the model. The largest, in August 1990, picks up the Iraq invasion of Kuwait. The others line up with OPEC meetings of various sorts. In one case, May 2004, the shock picks up an OPEC meeting and militant attacks that occurred in Saudi Arabia that month.[9] The newspaper index is forward-looking and occasionally

[9]A majority of the articles that month focused on the meeting. Of the 161 articles that contained the keyword OPEC, only 17 matched the keywords OPEC and terrorist and only 3 matched OPEC and militant.

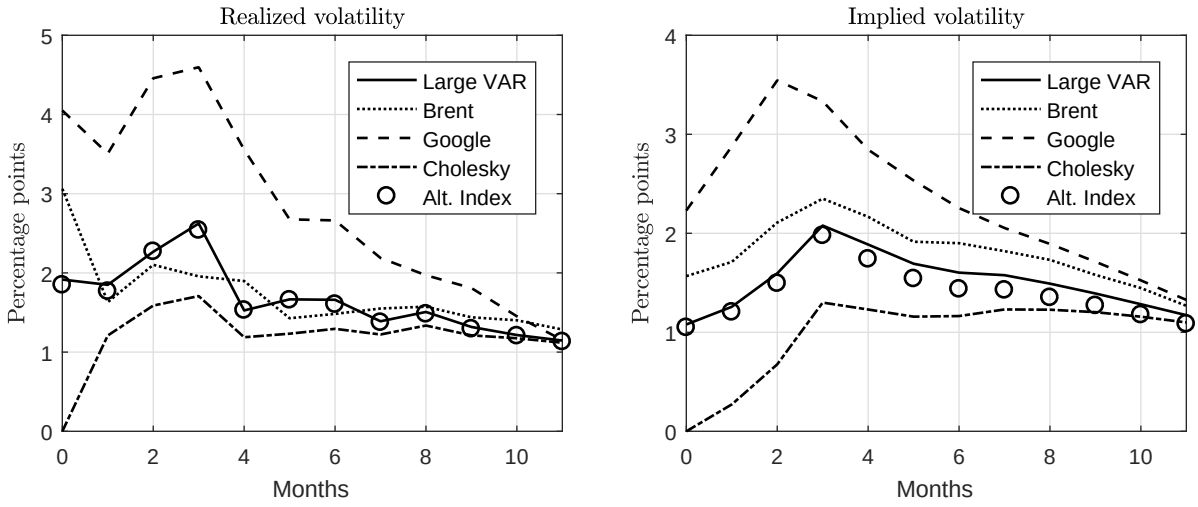

Figure 5: Generalized impulse response function for a shock to the OPEC newspaper index.

picks up unusual interest in events that occur later on. This is the case in October 2014, for example, when the index jumped in anticipation of a meeting in the following month. Beyond the top ten, many of the shocks also reflect meetings, with the next three largest shocks picking up an emergency ministers' meeting in April 1988, the OPEC meeting in October 2008 during the financial crisis and a regularly scheduled meeting in November 1997.[10] While many periods of unusual interest in OPEC line up with meetings, not every meeting generates the same interest levels. There were more than 90 official OPEC meetings from 1986 to 2016 but only a handful, those ranked from 2 to 6 in the table, resulted in shocks greater than 2 standard deviations in size.

Of the nine meetings listed in Table 2, there was one regularly scheduled bi-annual conference, four emergency meetings, one Ministerial Monitoring Committee (MMC) meeting, and three meetings that are more difficult to categorize. With regards to the latter, the first of these is the August 1986 meeting in Geneva, which ostensibly was a reconvening of the 78th OPEC meeting, a bi-annual conference meeting that occurred in June of that year. The second was an informal meeting held in Amsterdam in May 2004 that preceded an emergency meeting in early June 2004. The third was the meeting in November 2007, which was an OPEC "summit" held in Saudi Arabia.

One question of interest is why the amount of attention focused on OPEC was unusually high during those months listed in Table 2. To shed some light on this, I discuss some of the facts behind the three largest shocks associated with meetings: October 2014, October 2006 and March 1998. I also discuss the August 1986 meeting, given its unusual nature and the fact that it led to a production agreement. Discussion about the invasion of Kuwait, the largest shock, is omitted for brevity's sake given its well-known importance for the oil market.

The August 1986 meeting was officially a reconvening of an earlier meeting held in June 1986. It was preceded by a substantial decline in oil prices during 1986, due in part to a large Saudi

[10] Additional details about the shocks, including negative shocks, can be found in the appendix.

Arabian production increase from record lows in mid-1985. The August meeting attracted much attention due to the fact that it led to a surprise agreement on production levels.[11] WTI prices jumped almost 30 percent the week of the announcement. Unlike many of the other meetings listed in Table 2, in this case prices remained fairly stable for many months after the meeting.

Rank	Date	Events
1	August 1990	Iraq invades Kuwait
2	October 2014	Month prior to OPEC meeting (ordinary)
3	October 2006	OPEC meeting (emergency)
4	March 1998	OPEC meeting (emergency)
5	September 1993	OPEC meeting (emergency)
6	May 2004	Informal OPEC meeting
		Militant attacks in Saudi Arabia
7	September 1992	Ministerial Monitoring Committee meeting
8	August 1986	OPEC meeting, production agreement
9	November 2007	OPEC meeting (summit)
10	November 2001	OPEC meeting (emergency)

Table 2: This table lists the top ten shocks from the five variable VAR model with realized volatility.

Understanding the March 1998 meeting first requires discussing the November 1997 meeting, which was ranked as the 13th largest shock, as the two are connected. In early November 1997, Saudi Arabia's oil minister announced a desire to raise the OPEC production ceiling, interpreted by many as indicating Saudi Arabia's intention to increase production.[12] Indeed, Saudi production did increase in late 1997 and early 1998, as did output in the United Arab Emirates and Kuwait. In this particular meeting, there was an unexpected shift in OPEC policy that did not appear to be connected with a recent movement in oil prices.

The outcome of the November meeting eventually had implications for oil prices, which fell dramatically due to the increased supply, the return of Iraqi production and the impacts of the Asian financial crisis on demand. The March 1998 meeting attracted significant attention because there was internal disagreement between OPEC countries over how to respond to recent price declines. Reporting through most of March suggested no agreement would be reached. However, a surprise announcement regarding a production cut was made over the weekend of March 21 and 22.[13] Although prices stabilized temporarily, they eventually declined until they were close to $10 a barrel by year-end 1998.

No OPEC meeting was originally scheduled for October 2006 but an emergency meeting was held by the group on October 19 and 20. This meeting was spurred by a sharp decline in oil prices in the preceding months, with WTI having dropped more than 15 percent. This led to suggestions in early October of a possible meeting in the middle of month, although there were

[11] See "OPEC reaches tentative accord" (1986) [1].
[12] See Bahree and Fritsch (1997) [6] for more details.
[13] See Lucchetti and Ewing (1998) [23] for more details.

apparent disagreements over a potential production cut. An agreement was reached at the meeting, although there was some uncertainty reported among market participants about the deal.[14]

Finally, the events of October 2014 are not so distant in the past and perhaps require less discussion. Oil prices had fallen since mid-summer and by early October WTI prices were close to $90 a barrel. By the end of the month they were close to $80 a barrel. This is one case where the newspaper index exhibits some forward looking behavior, as the media was already writing an unusual number of articles mentioning OPEC despite the fact that its upcoming meeting was not until later in November. Similar to the October 2006 shock, media attention was focused on how OPEC might respond to the recent decline in oil prices.

A common feature across many of the meetings in table 2 is that they were preceded by a notable change in oil prices. Price declines occurred before the meetings in 1986, 1993, 1998, 2006 and in 2014, while increases preceded the meetings in May 2004 and November 2007. To illustrate this more concretely, I took the average (real) price of WTI in the month prior to the shock and compared it to the average price three months prior to the shock. So for example, comparing the average price in July 1986 to May 1986 I find that prices had declined by 29 percent. Prices declined by 10.1 percent, 23.1 percent, 15.8 percent and 10 percent for the 1993, 1998, 2006 and 2014 shocks, respectively. In the two periods where prices rose before the meetings, the increase was 5 percent for the 2004 event and more than 16 percent for the 2007 event. Newspaper articles suggest these price changes often helped generate increased interest in the OPEC event, as they led to discussion about how or if OPEC would respond.

One other interesting aspect of the October 2014 shock is that it is the result of events in the oil market that go beyond an OPEC meeting. The meeting in November 2014 was a regularly scheduled meeting and would have occurred regardless of those events but would probably not have garnered much, if any, unusual media interest had it not been for the price declines preceding it. This differs from the November 1997 shock, for example, which picked up a surprise policy change in production levels due to Saudi Arabia prodding. It also differs from the 1986 shock which, although driven by events in the oil market, was intimately connected with Saudi Arabia's decision to increase production in 1985 and 1986. Although not discussed in detail in this section, the October 2008 shock in this regard resembles the 2014 shock, as it was clearly being driven by events beyond OPEC, in this case the financial crisis.

4.3.2 DISCUSSION ON CAUSALITY

The models and methodology used here do not make any direct statements about causality between unexpected changes in the newspaper index and oil price volatility levels. It could be the case that some OPEC events are just unusually important, which drive media attention and volatility levels. Or it could be that other developments in the oil market or broader economy make an OPEC meeting or event more important and, thus, more widely watched while simultaneously leading to unexpectedly high levels of oil price volatility.

To the extent one can make such statements, the residual analysis suggests that both explanations play a role. For example, events such as the Iraq invasion of Kuwait are arguably exogenous events and garner unusual interest because of their importance. Likewise, although not the same as a war, the unexpected increase in the index that occurred in August 1986 was due to an event that was directly related to Saudi Arabia's prior decision to dramatically increase production levels. This has the sense of an OPEC policy shock, as it were. A similar argument could be made for the

[14]Please see Bahree (2006) [5], Foss (2006) [16] and Mouawad (2006) [25] for more details.

November 1997 shock.

On the other extreme, some unexpected increases in the newspaper index and price volatility appear to be driven by developments in the oil market or broader economy. For example, the large unexpected increase in the index in October 2014, reflecting an upcoming meeting, was obviously driven by oil price declines occurring around that time. This situation then raised questions about how OPEC might respond. A similar example would be the unexpected increase in October 2008, which was connected with developments in the broader economy. By October, the financial crisis had already affected oil prices, generated intense concern about future demand levels and lead to increased oil price volatility (both implied and realized). Here, developments in the economy that were also fueling volatility sparked interest in how OPEC might respond. In cases such as these, disentangling how much of the variation in oil price volatility is due to the underlying state of the world and how much is due to the OPEC events themselves appears to be a non-trivial task.

5 CONCLUSIONS

In this paper I introduce a monthly newspaper index that measures interest in OPEC events. I show that this index rises in response to important OPEC meetings, the Iraq invasion of Kuwait, and other events related to OPEC production levels, such as Saudi Arabia's decision to flood the market in the mid-1980s. The newspaper index is shown to be highly correlated with other measures of the public's interest in OPEC, such as Google Search Volume Index data. An advantage of the newspaper index is that it can be extended back much further in time than these alternatives.

I investigate the behavior of oil price volatility when there are unexpected changes in the newspaper index. There is very strong statistical evidence that unexpected increases in the index are associated with higher levels of oil price volatility, whether measured by implied volatility or realized volatility. This result is robust to several variations of the model, including using a different oil price in the model, using Google SVI data or employing an alternative formulation of the index. I find that many of the largest, unexpected increases in the index line up with specific events, primarily but not exclusively OPEC meetings. In some cases, interest levels in those events and price volatility appear to be endogenously responding to events in the oil market or broader economy.

Many possibilities present themselves for future research. One could envision doing more complicated textual analysis of the articles picked up in the search. Or extending this work to consider a broader set of events that have affected the oil market beyond those related to OPEC. The author is also currently investigating the potential usefulness of a higher frequency newspaper index for forecasting oil price volatility (or vice-versa).

REFERENCES

[1] Opec reaches tentative accord on iranian plan to limit output. *The New York Times*, August 5, 1986.

[2] Afkhami, M., Cormack, L., and Ghoddusi, H. (2017). Google search keywords that best predict energy price volatility. *Energy Economics*, 67:17–27.

[3] Amelier, C. and Darne, O. (2014). Volatility persistence in crude oil markets. *Energy Policy*, 65:729–742.

[4] Andrei, D. and Hassler, M. (2014). Investor attention and stock market volatility. *The Review of Financial Studies*, 28 (1):33–72.

[5] Bahree, B. Opec encounters hurdles in effort to trim output. *The Wall Street Journal*, October 6, 2006.

[6] Bahree, B. and Fritsch, P. Saudis signal desire to produce more oil as opec session looms. *The Wall Street Journal*, November 6, 1997.

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

[8] Baker, S. R. and Fradkin, A. (2017). The impact of unemployment insurance on job search: Evidence from google search data. *The Review of Economics and Statistics*, 99(5):756–768.

[9] Barber, B. M. and Odean, T. (2008). All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors. *Review of Financial Studies*, 21:785–818.

[10] Bernanke, B. S. (1983). Irreversibility, unvertainty, and cyclical investment. *The Quarterly Journal of Economics*, 98(1):85–106.

[11] Da, Z., Engleberg, J., and Gao, P. (2011). In search of attention. *Journal of Finance*, LXVI, no. 5:1461–1499.

[12] Demirer, R. and Kutan, A. M. (2010). The behavior of crude oil spot and futures prices around opec and spr announcements: An event study perspective. *Energy Economics*, 32:1467–1476.

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

[14] Engleberg, J. E. and Parsons, C. A. (2011). The causal impact of media in financial markets. *Journal of Finance*, LXVI (1):67–97.

[15] Ferderer, J. (1996). Oil price volatility and the macroeconomy. *Journal of Macroeconomics*, 18(1):1–26.

[16] Foss, B. Opec move leaves doubts. *Houston Chronicle*, October 21, 2006.

[17] Grossman, V., Mack, A., and Martinez-Garcia, E. (2014). A new database of global economic indicators. *Journal of Economic and Social Measurement*, 39(3):163–97.

[18] Guo, H. and Kliesen, K. L. (2005). Oil price volatility and u.s. macroeconomic activity. *Federal Reserve Bank of St. Louis Review*, 87(6):669–683.

[19] Horan, S. M., Peterson, J. H., and Mahar, J. (2004). Implied volatility of oil futures options surrounding opec meetings. *The Energy Journal*, 25(3):103–125.

[20] Jo, S. (2013). The effects of oil price uncertainty on the macroeconomy. forthcoming, *Journal of Money, Credit and Banking*.

[21] Kellogg, R. (2014). The effect of uncertainty on investment: Evidence from texas oil drilling. *American Economic Review*, 104(6):1698–1734.

[22] Kilian, L. (2009). Not all oil price shocks are alike: Disentangling demand and supply shocks in the crude oil market. *American Economic Review*, 99(3):1053–69.

[23] Lucchetti, A. and Ewing, T. Oil prices rocket 13%, stir commodity sector. *The Wall Street Journal*, March 24, 1998.

[24] Mensi, W., Hammoudeh, S., and Yoon, S.-M. (2014). How do opec news and structural breaks impact returns and volatility in crude oil markets? further evidence from a long memory process. *Energy Economics*, 42:343–354.

[25] Mouawad, J. From opec, conflicting talk of output cut. *The New York Times*, October 6, 2006.

[26] Pesaran, H. H. and Shin, Y. (1998). Generalized impulse response analysis in linear multivariate models. *Economics Letters*, 58:17–29.

[27] Robe, M. A. and Wallen, J. (2016). Fundamentals, derivates market information, and oil price volatility. *Journal of Futures Markets*, 36(4):317–344.

[28] Schmidbauer, H. and Rosch, A. (2012). Opec news announcements: Effects on oil price expectation and volatility. *Energy Economics*, 34(5):1656–1663.

[29] Tetlock, P. C. (2007). Giving content to investor sentiment: The role of media in the stock market. *The Journal of Finance*, LXII (3):1139–1168.

[30] Yuan, Y. (2015). Market-wide attention, trading and stock returns. *Journal of Financial Economics*, 116:548–564.

NOT FOR PUBLICATION

A APPENDIX

A.1 FACTIVA DATA The Factiva searches make use of the Search Builder command. I perform a "Free Text Search" using the keyword OPEC. The duplicates toggle is set to "Identical." The exact sources are The New York Times, print edition (newspaper icon); The Wall Street Journal, print edition (newspaper icon); the Financial Times (Available through Third Party Subscription Services), print edition (newspaper icon); and the Houston Chronicle, print edition (newspaper icon). The toggles "Search Source Name/Alias only" and "Exclude Discontinued Sources" are checked. In some months, Factiva erroneously reports duplicates of an article as being distinct from each other. To deal with this, the articles were checked by hand and those duplicates removed from the counts. The results are not very different whether or not these duplicates are removed. Figure 6 plots the standardized time series for each of the four newspaper articles.

To generate the estimate for the total volume of news articles each month I repeat the search but replace the keyword "OPEC" with "the or a or an." Given the grammatical structure of English, which requires these three words to appear frequently in sentences, this should capture most, if not all, of the articles written in the newspapers considered.

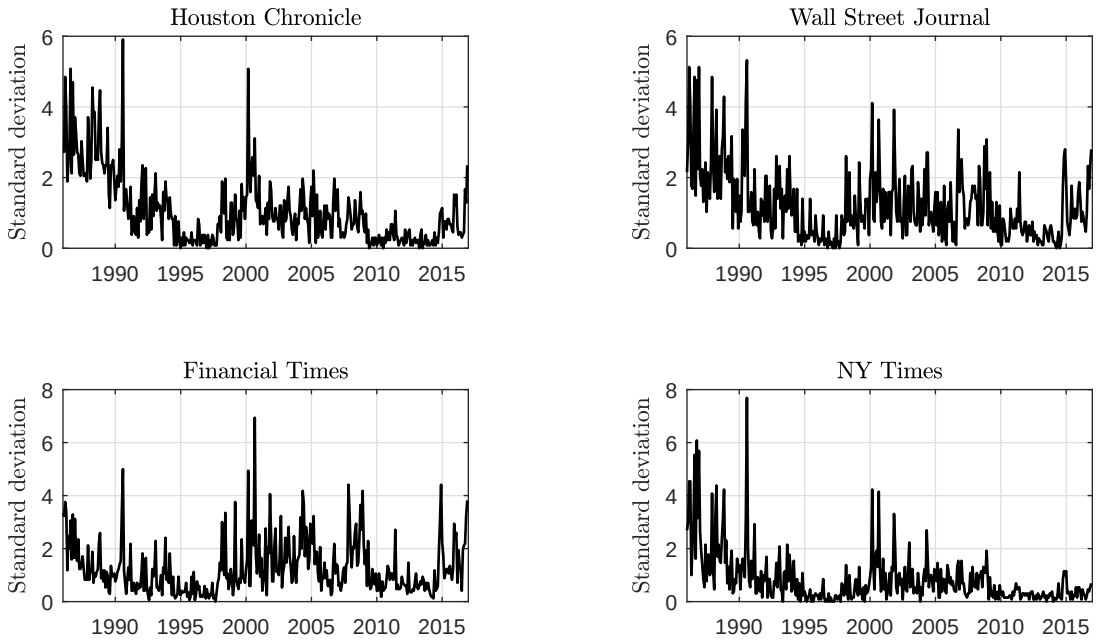

Figure 6: Standardized time series for the individual newspaper series.

A.2 GOOGLE DATA AND RESULTS The Google search volume data is produced in the following manner. I use the Google Trends web site and get the data for OPEC as a topic, not a keyword. The geographic area is the entire world. The sample runs from 2004 to 2016 and Google Trends returns the data in monthly frequency.

Table 3 presents the correlations between the Google data, the OPEC newspaper index and the individual counts. The Google data has a fairly high correlation with the newspaper index and the individual counts.

Residual analysis was also done with the model using Google SVI data. As the sample starts in 2004, only a subset of the events that might be picked up by the newspaper index are covered with the SVI data. Table 4 lists the top 10 shocks. Many of the events that lay behind these unexpected increases in search activity on OPEC were also picked out as unusual events in the baseline model.

Google SVI	1.00					
Financial Times	0.75	1.00				
Houston Chronicle	0.72	0.76	1.00			
New York Times	0.74	0.70	0.73	1.00		
Wall Street Journal	0.62	0.79	0.78	0.70	1.00	
OPEC index	0.78	0.93	0.89	0.83	0.92	1.00

Table 3: Correlations between Google SVI, the OPEC index and the individual newspaper counts. The sample runs from Jan. 2004 to Dec. 2016.

Rank	Date	Events
1	November 2014	OPEC meeting
2	November 2007	OPEC summit
3	October 2008	OPEC meeting (emergency)
4	October 2006	OPEC meeting (emergency)
5	May 2004	Informal OPEC meeting
		Militant attacks in Saudi Arabia
6	August 2005	Oil price surge, month prior to meeting
7	December 2015	OPEC meeting
8	May 2008	Militant attacks in Nigeria
9	June 2004	OPEC meeting (emergency)
10	September 2016	OPEC meeting (emergency)

Table 4: This table lists the top ten shocks for the VAR model that uses Google data in place of the OPEC newspaper index.

A.3 ADDITIONAL RESULTS Table 5 lists the top 20 positive shocks from the baseline model. In general, the shocks ranked 11 to 20 occur either one month prior to or during the month of an OPEC meeting. The 14th largest shock, which is in April 1997, does not pick up any particular event. Instead, it is statistical in nature and driven by the fact that in the prior month there was an exceptionally low amount of articles written on OPEC. As the index is persistent in nature, the increase that occurred in April appears as a large positive shock to the index. The 16th largest shock picks up the Venezuela coup attempt in April 2002.

Rank	Date	Events
1	August 1990	Iraq invades Kuwait
2	October 2014	Month prior to OPEC meeting (ordinary)
3	October 2006	OPEC meeting (emergency)
4	March 1998	OPEC meeting (emergency)
5	September 1993	OPEC meeting (emergency)
6	May 2004	Informal OPEC meeting
		Militant attacks in Saudi Arabia
7	September 1992	Ministerial Monitoring Committee meeting
8	August 1986	OPEC meeting, production agreement
9	November 2007	OPEC meeting (summit)
10	November 2001	OPEC meeting (emergency)
11	April 1988	Ministerial meeting (emergency)
12	October 2008	OPEC meeting (emergency)
13	November 1997	OPEC meeting (ordinary)
14	April 1997	No event (index jumps from low in March 1997)
15	February 2000	Month prior to OPEC meeting
16	April 2002	Venezuela coup
17	August 2002	Month prior to OPEC meeting
18	March 1991	Ministerial Monitoring Committee meeting
19	June 2000	OPEC meeting (emergency)
20	July 2001	OPEC conference

Table 5: This table lists the top twenty shocks from the five variable VAR model with realized volatility.

Table 6 lists the ten most negative shocks in the model. By construction, these occur during months when there were unexpectedly low article counts. In some cases, this is literally because the number of articles written about OPEC was very low relatively to the norm. These shocks are clustered in the mid-1990s, a period of relatively stable oil prices. Several of the shocks occur in a month following an OPEC meeting. This occurs for two reasons. First, in general there are more articles written about OPEC during months with meetings. Second, the index is quite persistent, on average, so if following a meeting there is a sharp drop in articles, this will show up as a negative shock in the model.

Rank	Date	Events
1	March 1997	Only two articles published
2	July 1994	Month following an OPEC meeting
3	February 1996	Only six articles published
4	July 1996	Month following an OPEC meeting
5	September 1994	Only eight articles published
6	December 1996	Only five articles published
7	June 1992	Month following an OPEC meeting
8	April 2004	Month following an OPEC meeting
9	September 1997	Only four articles published
10	February 2015	Sharp drop in articles published

Table 6: This table lists the most negative shocks from the five variable VAR model with realized volatility.