NCER Working Paper Series

The impact of information flow and trading activity on gold and oil futures volatility

A E Clements
N Todorova

Working Paper #102
June 2014
The impact of information flow and trading activity on gold and oil futures volatility

Adam E. Clements∗

School of Economics and Finance, Queensland University of Technology

Neda Todorova†

Department of Accounting, Finance and Economics, Griffith University

Abstract

There is a long history of research into the impact of trading activity and information on financial market volatility. Based on 10 years of unique data on news items relating to gold and crude oil broadcast over the Reuters network, this study has two objectives. It investigates the impact of shocks in trading activity and traders positions which are unrelated to information flows on realized volatility. Additionally, the extent to which the volume of the information flow as well as the sentiment inherent in the news affects volatility is also examined. Both the sentiment and rate of news flow are found to influence volatility, with unexpected positive shocks to the rate of news arrival, and negative shocks to the sentiment of news flow exhibiting the largest impacts. While volatility is also related to measures of trading activity, their influence decreases after news is accounted for indicating that a non-negligible component of trading is in response to public news flow. After controlling for the level of trading activity and news flow, the net positions of the various types of traders play no role, implying that no single group of traders lead to these markets being more volatile.

Keywords: Information flow; Volatility; Oil futures; Gold futures; Trading activity.

JEL Classification Numbers: C22; G10; G13; G14.

∗a.clements@qut.edu.au
†n.todorova@griffith.edu.au
1 Introduction

There is a long history of research on the impact of trading activity and information on financial market volatility. Based on an unique data set of news items on gold and crude oil broadcast over the Reuters network, this study has two fundamental objectives. First, it investigates the impact of shocks in overall trading activity and traders positions unrelated to information flows on gold and oil futures volatility. Second, it analyses the extent to which the volume of the information flow as well as the sentiment inherent in the news affects the realized volatility of these markets.

Investors’ interest in commodities, such as crude oil and gold has risen dramatically in the last decade. In the oil market the term “paper barrels” has become increasingly popular, reflecting growth in trading in oil markets via derivatives without any interest in the physical commodity itself (Yergin, 2012) while gold has established its status of a safe haven (Baur and McDermott, 2010). Thus, it is important to examine whether the volatility in these markets is influenced by trading activity which cannot be explained by public information flow, and if activities of individual trader groups make futures markets more volatile. In this paper, we extend extant studies by using the weekly reports of the Commodity Futures Trading Commission (CFTC) on the outstanding futures positions by type of trader.\(^1\) While the majority of studies using CFTC data focus on the contemporaneous explanatory power or predictive ability of the Commitment of Trader (COT) positions regarding futures or spot returns (e.g. Wang, 2003, Sanders, Boris, and Manfredo, 2004, Schwarz, 2012, Chen and Maher, 2013), the relation between trader positions and volatility has been addressed from a number of different perspectives. Chang, Chou, and Nelling (2000) study the relation between stock market volatility and the demand for hedging in S&P 500 stock index futures contracts. Pan, Liu, and Roth (2003) examine how volatility and the futures risk premia affect trading demands for hedging and speculation. Roethig and Chiarella (2007) examine the nonlinearities in the response of speculators’ trading activity to price changes in live cattle, corn, and lean hog futures markets. Tornell and Yuan (2012) analyze the relationship between futures trading activities of speculators

---

\(^1\)The CFTC classifies reportable positions as either commercial or noncommercial based on whether a trader holds a reportable position. Traders taking derivatives positions to hedge specific risks are commonly regarded as hedgers. Noncommercial traders who trade futures for reasons other than hedging are seen as speculators. Traders with nonreportable positions are referred to as small traders.
and hedgers and potential movements of major spot exchange rates.


The novelty of the current study is that it explicitly accounts for the impact of information flow on popular trading activity measures. Existing empirical research on the relation between volatility and news focuses mostly on the impact of specific news events such as macroeconomic announcements, political interventions or earnings announcements (e.g. Ederington and Lee, 1993, Christie-David and Chaudhry, 1999, Hautsch, Hess, and Veredas, 2011). There are few studies investigating the role of news sentiment or number of news items and these studies address markets for assets other than commodity futures. Kalev et al. (2004) study the information-volatility relation for Australian equities and proxy information flow by the number of all news announcements made on individual companies. Gross-Klussmann and Hautsch (2011) use news data for individual equities traded at the London Stock Exchange sampled at high frequency to examine how news arrivals influence trading activity in individual stocks. Ho, Shi, and Zhang (2013) examine the relationship between the volatility of Dow Jones 65 stocks and public news sentiment based on firm-specific and macroeconomic news announcements and their sentiment scores. Riordan et al. (2013) focus on the impact of newswire messages on intraday price discovery, liquidity, and trading intensity of Canadian stocks. Using news headlines relating to constituents of the S&P 500, Smales (2014) examines the relationship between aggregate news sentiment and changes in the implied volatility index VIX.

This paper makes two major contributions. First, the methodology of Bessembinder and Seguin (1992), Wang (2002a) and Daigler and Wiley (1999) who decompose popu-
lar measures of trading activity into expected and unexpected components is extended. The current analysis is the first to address the impact of trading activity on realized volatility by estimating trading activity shocks which are unrelated to information flow. Based on an advanced volatility proxy established with intraday data, we find a negative and insignificant relation between shocks in net speculators’ positions in the oil futures market and a negative and only weakly significant relation between volatility and the corresponding positions in the gold futures market after controlling for the information flow. The results are robust to the choice of volatility proxy and indicate that the level of activity of particular types of traders does not affect significantly the level of volatility. Our findings are in line with Wang (2002b), Bryant, Bessler, and Haigh (2006), and Bohl and Stephan (2013) and contribute to the ongoing controversial discussion whether speculative positions destabilize the market. Second, this is also the first study which analyses the impact of the volume of the information flow and its sentiment on volatility of gold and oil futures. Additionally, the number of news items and their sentiment is decomposed into expected and unexpected components which demonstrates the relative importance of shocks. When the direction of the shocks is taken into account, it becomes obvious that negative shocks in news sentiment and positive shocks in the volume of news have the strongest impact.

The paper is arranged as follows. The next sections outline the data and the methodology. Subsequent sections present the empirical results and concluding remarks, respectively.

2 Data

2.1 Intraday futures prices and volatility estimation

This study uses several different data types available for the period of January 2003 to October 2012. First, realized volatility is estimated using 5-minutes returns for futures on the crude sweet oil Western Texas Intermediate (WTI) traded on NYMEX and gold futures traded on COMEX. The contract size is 1,000 barrels for oil futures and 100 troy ounces for gold futures, respectively. The data originate from the exchange operated electronic trading platform Globex and are obtained from the Thomson Reuters Tick
History database at the Securities Industries Research Centre of Asia Pacific (Sirca). This study rolls the nearest month contract to the next most liquid month when the daily volume of the current contract is exceeded.

To estimate the daily quadratic variation using intraday data, the realized variance as proposed by Andersen and Bollerslev (1998) is employed. If \( r_{\tau,i} \) denotes the \( i \)th intraday return on day \( \tau \), the realized variance on a day \( \tau \) is estimated by finding the total of the squared intraday returns:\(^2\)

\[
RV^2_{\tau} = \sum_{i=1}^{I} r^2_{\tau,i}.
\]

(1)

Futures are traded, except for short interruptions, almost around the clock on weekdays. Following the contract specifications of the gold and oil futures under consideration, a trading day is defined as the interval from 6 pm (ET) of one day to 5:59 pm of the next day. Intraday prices of weekend periods (between Friday 6 pm and Sunday 6 pm) and days corresponding to US public holidays are removed.

To match the COT data, which is only publicly available at a weekly frequency, daily realized variances are aggregated to weekly realized volatilities as the annualized square root of the average daily variance estimates from Wednesday of week \( t-1 \) to Tuesday of week \( t \),

\[
RV_t = \sqrt{252 \cdot \frac{1}{N} \sum_{j=1}^{N} RV^2_{t-j}}.
\]

(2)

with \( N = 5 \) for weeks with no public holidays and \( N \) corresponding to the actual number of trading days in the week otherwise.

Existing studies on the linkage between trading activity and volatility document that the results are often very sensitive to the volatility estimate used (e.g. Pan, Liu, and Roth, 2003, Luu and Martens, 2003), especially when findings based on advanced

\(^2\)There is a plethora of more efficient and cumbersome ways to account for market microstructure effects for the purpose of estimating realized volatility. Liu, Patton, and Sheppard (2013) compare the estimation and forecasting performance of around 400 different realized measures and document that whether the five-minute realized volatility can be consistently outperformed depends on the utilized benchmark. However, as Bollerslev, Tauchen, and Zhou (2009) point out, “the simple-to-implement realized volatility estimator based on the summation of (not too finely sampled) high-frequency squared returns [..] remains the dominant method in practical applications”.

estimators from the family of the realized measures are compared with studies using volatility proxies based on more noisy data such as daily closing prices. To make the findings comparable with previous studies (e.g. Bessembinder and Seguin, 1992, Wang, 2002a) and to establish whether results are sensitive to the use of realized volatility, the daily Parkinson (1980) range,

$$PK^2_{\tau} = \frac{(H_{\tau} - L_{\tau})^2}{4 \ln 2},$$

(3)

where $H_{\tau}$ and $L_{\tau}$ denote the log of highest and lowest price of day $\tau$ is also employed. Following equation (2), the final volatility estimate is the square root of the annualized daily averages from Wednesday of week $t-1$ to Tuesday of week $t$, respectively.

### 2.2 Trading activity data

Trading volume data originates from Datastream and comprises the daily trading volumes of the gold and oil futures transactions taken place through the Globex, ClearPort and the CME Open Outcry system. Furthermore, this paper analyzes total open interest and trader position data from the COT reports of the closing positions aggregated for all outstanding contracts categorized by trader type. Data are collected by the CFTC each Tuesday, and reflect positions on that day. As mentioned in the introduction, the CFTC classifies reportable positions as either commercial or noncommercial based on whether a trader holds a reportable position. The reportable level in terms of number of futures contracts is 350 contracts for WTI futures and 200 contracts for gold futures.

### 2.3 Information flow data

To capture news flow, pre-processed news data from the Thomson Reuters News Analytics database is used. The text of news items broadcast over the Reuters network is analysed using linguistic pattern recognition algorithms. The analysis produces a number of characteristics relating to each news item including relevance to the specific firm, sentiment and novelty. Sentiment for each news item is coded +1, 0, -1 for positive, neutral and negative tones respectively. Here, we take all news articles denoted as articles, representing fresh stories consisting of a headline and body text. Appendts to previous articles,
and alerts with no text body are ignored. News items relating to the topic codes GOL (Gold) and CRU (Crude Oil) are collected from the commodities files for each day during the sample period with a number of measurements relating to news flow taken. The total number of news items (denoted below as $N$) across is recorded reflecting the volume of information flow. The average sentiment (denoted below as $S$) across the news items for each day is taken to capture the overall tone of news flow.

### 2.4 Summary statistics

As this study is conducted at a weekly frequency, daily volatilities, trading volume, news sentiment and number of news items are aggregated to weekly averages from the preceding Wednesday to the following Tuesday. Tables 1 and 2 present summary statistics for weekly realized and range-based volatilities, returns, news variables (sentiment and number of news), overall trading activity (open interest and trading volume) and net positions by type of trader. The average daily number of news at a weekly frequency ($N$) is in units of 100 whereas the trading volume ($VOL$) and the total open interest ($OI$) are in units of 100,000 futures contracts.

Both oil and gold futures exhibit positive average weekly returns with the oil futures prices exhibiting almost double the volatility of gold futures. On average, weekly sentiment is positive with the news items about crude oil having a slightly higher sentiment than the news about gold. The number of news items relating to the oil market on average per day is threefold the number of news associated with the gold market. Figure 1 provides a general view of the dynamics of realized volatility and the course of news sentiment and number of news items over time. The results in the upper panels of Tables 1 and 2 also include the Augmented Dickey-Fuller (ADF) test statistics for the presence of a unit root in the corresponding time series. The null hypothesis of an existing unit root can be conclusively rejected for all series but the total open interest for the sample period under consideration.

The lower panels of Tables 1 and 2 present the contemporaneous correlations between all considered variables. Realized volatility is positively correlated with the number of news items indicating that increased information arrivals are associated with more volatile markets, as expected. This correlation is stronger for gold than for oil futures. Moreover,
realized volatility is negatively correlated with the sentiment inherent in the news flow confirming that decreasing news sentiment results in higher volatility. The net positions of speculators and small traders (hedgers) in the oil market are negatively (positively) correlated with realized volatility, news sentiment and number of news items while the net positions of speculators exhibit weak correlation with gold futures volatility.

3 Methodology

To understand the overall relation between information flow and volatility, a regression model of the following form for the volatility $\sigma_t$ is estimated,

$$
\sigma_t = \alpha_0 + \sum_{i=1}^{m} \alpha_i \sigma_{t-i} + \beta_1 S_t + \beta_2 N_t + \epsilon_t.
$$

where $RV_t$ represents $\sigma_t$ in this first case. Lagged volatilities are included to account for the effect of volatility persistence. Additionally, the weekly news sentiment $S_t$ and the volume of the information flow $N_t$ are decomposed into expected ($S_t^E, N_t^E$) and unexpected components ($S_t^U, N_t^U$) in order to isolate the impact of surprises in news flow within this relation,

$$
\sigma_t = \alpha_0 + \sum_{i=1}^{m} \alpha_i \sigma_{t-i} + \beta_1 S_t^E + \beta_2 S_t^U + \beta_3 N_t^E + \beta_4 N_t^U + \epsilon_t.
$$

The expected component of each news measure is the fitted value from an ARIMA(p,d,q) model, while the unexpected component is the actual sentiment or number of news less the expected component. The number of lags is chosen based on the Akaike information criterion and regressions are run with various lag lengths to ensure robustness. Since both the weekly news sentiment and volume of information flow series are stationary, a ARIMA(3,0,0) model is used for decomposition. The coefficients for lags higher than three prove to be insignificant.

It is well known that financial market volatility responds asymmetrically to positive and negative news. Therefore, to examine potential asymmetric responses to shocks in information flow, the residual series $S_t^U$ and $N_t^U$ is split into two series reflecting only positive ($S_{t,+}^U$ and $N_{t,+}^U$) and only negative shocks ($S_{t,-}^U$ and $N_{t,-}^U$), respectively. This
approach provides a complete decomposition of $S_t^U$ and $N_t^U$ in that $S_t^U = S_{t,+}^U + S_{t,-}^U$ and $N_t^U = N_{t,+}^U + N_{t,-}^U$. The corresponding regression model takes the form,

$$\sigma_t = \alpha_0 + \sum_{i=1}^{m} \alpha_i \sigma_{t-i} + \beta_1 S_{t,+}^U + \beta_2 S_{t,-}^U + \beta_3 N_{t,+}^U + \beta_4 N_{t,-}^U + \beta_5 N_{t,+}^U + \beta_6 N_{t,-}^U + \epsilon_t. \quad (6)$$

Models (4) to (6) involve solely contemporaneous values of the information flow variables since preliminary analysis has shown that lagged values remain consistently insignificant.

To examine the impact of shocks in trading activity on volatility, a regression model of the following form is estimated,

$$\sigma_t = \alpha_0 + \sum_{i=1}^{m} \alpha_i \sigma_{t-i} + \beta_1 V_{LU}^U + \beta_2 O_{LU}^U + \beta_3 N_{LU,j}^U + \epsilon_t \quad (7)$$

where $V_{LU}^U$ and $O_{LU}^U$ are the unexpected trading volume and total open interest, respectively, and $N_{LU,j}^U$ is the unexpected net trader position. $j$ represents trader type, speculators, hedgers, and small traders. The net positions on Tuesday of week $t$ are established by

$$N_{LU,j}^U = L_{LU,j}^U - S_{LU,j}^U \cdot O_{LU}^U, \quad (8)$$

with $L_{LU,j}^U$ ($S_{LU,j}^U$) being the number of outstanding long (short) positions by trader type. Similar to Wang (2002a), due to the high correlations between net positions of the different trader types (-0.978 for oil and -0.921 for gold between net positions of speculators and hedgers), equation (7) is estimated separately for each trader type.

To establish the components of the trading activity variables which are not related to news impacts, $V_{LU}$, $O_{LU}$ and $N_{LU,j}^U$, are decomposed into expected and unexpected components by using ARIMAX(p,d,q) model including the information flow variables $S$ and $N$ as exogenous variables. The expected component is the fitted value from the ARIMAX model, while the unexpected series is the observed residual. The existence of a unit root has implications for splitting a variable into elements related to its own history as well as the information flow and unexpected elements. A stationary variable is decomposed with an ARIMAX(3,0,0) model including the contemporaneous values and

---

3The impact of news and trading volume shocks are analysed via separate regressions to ensure a clearly-arranged presentation. Including news sentiment and number of news in equation (7) does not alter the results as the series of unexpected trading volume, open interest and net positions of traders are generated by filtering out the informational content of the news items.
the first lags of the news sentiment and number of news items as exogenous variables. A variable with a unit root requires an ARIMAX(3,1,0) model with the same exogenous variables. In the majority of the cases, the contemporaneous values of the news sentiment and number of news are significant while the results for the coefficients of their first lags are more mixed. Even if lags beyond the first one remain insignificant, the analysis is run with various specifications to ensure robustness (results not reported).

Additionally, the regressions (7) for the individual trader types are rerun by using unexpected trading volume, total open interest and net trader positions generated without taking news sentiment and number of news into account. This ensures comparability with existing studies which do not explicitly consider the information flow.

4 Results

4.1 Volatility and news

Table 3 presents the regression results for equations (4) to (6) involving weekly realized volatility and weekly averages of news sentiment and number of news items.\(^4\) To begin, realized volatility of oil and gold futures exhibits the common feature of strong persistence. In both markets, volatility is significantly related to both news sentiment and number of news items. News sentiment is negatively related to volatility, as expected, indicating that decreasing sentiment makes these commodity markets more volatile. The number of news items is found to be positively associated with volatility meaning that an increasing volume of information arrivals drives volatility higher.

When news sentiment is decomposed into its expected and unexpected components, it becomes obvious that the expected sentiment is significant only at the 10% level while the unexpected component is significant at the 1% level with a magnitude in absolute terms around two times greater than the expected component. Furthermore, the negative coefficients of the unexpected sentiment in equation (5) confirm that especially shocks in news sentiment lead to increasing volatility. Splitting the sentiment shocks into two series depending on the direction of the shock (equation (6)) sheds further light on the news-volatility relation. The negative relation between unexpected news sentiment and

\(^4\)The results obtained with the daily range (3) are similar and not reported for the sake of brevity.
realized volatility can be conclusively ascribed to negative shocks. In other words, volatility increases significantly when the news sentiment turns out to be lower than expected, while if the news’ tone is more positive than expected, no significant impact on volatility can be documented. In contrast to sentiment, both expected and unexpected volume of information flow have a positive and significant impact on realized volatility indicating that the volume of news about gold and oil markets is generally strongly related to the volatility. The volume of information flow affects the general level of volatility, regardless of whether the news are positive or negative, expected or unexpected. Similar to the overall sentiment inherent in news, the unexpected components are of greater importance than their expected counterparts, with estimated coefficients of 0.108 and 0.064 for gold, and 0.046 and 0.026 for oil (all significant at the 1% level), respectively. Analysing the direction of the shocks uncovers that the unexpected rate of news arrival makes gold futures prices more volatile regardless whether the information volume is more or less than expected. However, a positive shock appears to have a higher impact (coefficient of 0.133) than a negative deviation (0.083). In contrast, only a larger amount of news leads to significantly higher volatility in the oil market. Informally, these results suggest that the presumption of “no news is good news” is more accurate for the oil rather than the gold market.

4.2 Volatility and shocks in trading activity

Tables 4 and 5 present the regression results for equation (7) when realized volatility is used as a proxy of the variation in the futures oil and gold returns. The upper panel in each table shows results when unexpected trading activity is based only on its own lags. The lower panel shows the corresponding results when both its own history as well as news flow are taken into consideration. A comparison of the regression results in both panels shows whether the impact of trading activity on volatility changes when news flow variables are taken into account for extracting the expected components. Doing so highlights the impact of trading activity that is unrelated to news flow. To visualise how the estimates in (7) behave when considering shocks of the individual trading activity measures, regression results are first presented for the case when (7) includes lagged volatilities only (Model 1). Then, the impact of unexpected trading volume and open
interest are assessed separately (Model 2 and 3, respectively). Models 4 to 6 include the unexpected net positions by trader type (speculators, hedgers and small traders, respectively) along with the unexpected trading volume and open interest. The latter three regressions allow for analysing the relevance of unexpected trading activity of specific types of traders after controlling for the overall trading activity.\footnote{We focus on the unexpected series as our preliminary analysis confirmed the findings of previous studies that these series have a higher impact than the expected components of trading activity in explaining volatility.}

One important result here is that the impact of all trading activity measures on volatility decreases once news flow is accounted for. This general result implies that a non-trivial component of trading is in response to public news flow. Shocks in trading volume appear to be positively related to volatility and strongly significant for both markets, regardless of whether news flow is taken into account. This indicates that the results of Bessembinder and Seguin (1992) and Wang (2002a) hold after filtering out the information content of the contemporaneous information flow. In line with these studies, open interest consistently has a negative coefficient as it is related to the number of traders or amount of capital in a market. Since these factors enhance market depth, deeper markets are associated with lower volatility shocks (Bessembinder and Seguin, 1992). However, while the unexpected open interest remains significant after filtering out the information inherent in news about crude oil, it becomes insignificant for realized volatility in the gold futures market.

Models 4-6 highlight the impact of trader positions. The net positions of small traders are insignificant in all cases, in line with the expectation that the trading activity exhibited by this trader group is least informative due to its heterogeneity. While the coefficients of the net positions of both hedgers and speculators are significant for the gold market when decomposition is conducted solely based on the own history, the only trader type in the oil market with a significant impact on volatility is the group of hedgers, and is only significant at 10%. Once news flow is taken into account, the impact of the net trader positions are insignificant with the exception of the speculators’ positions in the gold futures market (significant at the 10% level). It seems as though trading positions react to information in news flow, once this is accounted for, there is no residual impact on volatility.
In terms of direction, volatility appears to be negatively associated with shocks in net positions of speculators and small traders, and positively related to shocks in net positions of hedgers in both the oil and gold market, no matter whether the information flow is taken into account or not. This finding is at odds with the study of Wang (2002a) on the foreign exchange market and provide no evidence that changes in speculative positions destabilize the market. What can be concluded is that speculators seem to possess some private information in the oil and gold markets, whereas hedgers are less informed and their trading activities make the markets more volatile. However, the impact of both trader types on volatility is statistically insignificant. This is in line with Wang (2002b), who split trading demand by type of trader in the S&P 500 index futures market into expected and unexpected components and find that volatility is negatively related to speculative demand shocks and positively related to hedging demand shocks. Our findings also support the conclusions of Bryant, Bessler, and Haigh (2006) and Bohl and Stephan (2013) who, using a different methodology, find no evidence that the level of activity of particular types of traders influences the level of volatility.

The information content of the trading activity variables can be further assessed by taking the increments in the adjusted $R^2$ of the conducted regressions. It becomes obvious that the most pronounced increase is driven by the shocks in trading volume while net trading positions and especially open interest in isolation have negligible contributions, especially once the effect of the information flow has been filtered out from the shocks in these series. This is an intriguing observation which is at odds with Wang (2002a) who documents pronounced increases in $R^2$ in the foreign exchange market by considering net traders positions. However, Wang (2002a) employs daily returns or daily high-low ranges to estimate volatility, documenting much lower volatility persistence. Our conclusions are based on a more advanced intraday data based volatility proxy. To assess whether the choice of the volatility proxy affects our overall findings and can explain why our results deviate from those of Wang (2002a), the models 1 to 6 in Tables 4 and 5 are re-estimated using the daily Parkinson range.

The regression results obtained with the high-low range are presented in tables 6 and 7. Overall, the results do not appear to be sensitive to the choice of volatility proxy. Shocks in trading volume remain strongly significant even after controlling for the news flow with
smaller decreases in magnitude than in the RV case. Shocks in total open interest and net positions of the different trader types, which are partially significant if only own lags are taken into account, become consistently insignificant when the news sentiment and number of news are included. Interestingly, the volatility persistence as captured by the sum of lagged volatilities remains of a similar magnitude as in the case of RV which is overall much higher than in the studies of Wang (2002a) and Wang (2002b). The main difference between running regression (7) with realized volatility and the daily range is uncovered by comparing pairwise the explanatory power of the models in tables 4 to 7. The adjusted $R^2$ remain quite high even when moving from RV to daily range but the overall increase driven by including trading activity shocks is consistently more pronounced for the daily range. Including unexpected trading volume enhances the adjusted $R^2$ by almost 0.08 for gold when the news flow is not taken into account and 0.067, otherwise. The corresponding increases in $R^2$ with the oil price daily range are 0.038 and 0.032, respectively (Table 6) which is also higher than in the case of realized volatility (Table 4). Consistent with earlier results, shocks in total open interest and net positions of the different trader types appear not to have a significant impact on volatility.

To sum up, our findings indicate that the volume and tone of the information flow have a pronounced impact on volatility. Once they are filtered out of popular trading volatility measures, only shocks in trading volume remain consistently significant in explaining weekly futures volatility in the oil and gold futures markets. Without controlling for the information flow, the net traders positions exhibit a degree of significance which, however, disappears if news is controlled for. Additionally, adding net positions does not yield higher adjusted $R^2$. These results are at odds with Wang (2002a), who documents that volatility is positively associated with shocks in net positions of speculators and negatively related to shocks in net positions of hedgers in the foreign exchange market, and Chang, Pinegar, and Schachter (1997) who observe a positive relation between speculative trading volume and price volatility in the S&P 500 index, Treasury bonds, gold, corn, and soybean futures markets. These findings are more in line with Wang (2002b) who finds that volatility is negatively related to speculative demand shocks and positively related to hedging demand shocks in the S&P 500 index futures market. This implies that much of the impact of traders activity on volatility is due to public news flow. Additionally, our re-
sults support the conclusions of Bryant, Bessler, and Haigh (2006) and Bohl and Stephan (2013) who find no evidence that the level of activity of particular types of traders affects significantly the level of volatility. Based on a subsample analysis, Mutafoglu, Tokat, and Tokat (2012) focus on the interrelation between the trader positions in gold, silver and platinum futures markets and the related spot market return and document that the causal link from traders positions to market returns has weakened over time. Our findings suggest that this development is evident in the second moment of returns as well.

5 Conclusion

There is a long history of research considering the empirical explanations of asset return volatility. Broadly speaking, two main determinants have been considered, measure of trading activity and information flows. In the context of gold and oil futures markets, this paper has considered the role of shocks in overall trading activity and traders positions, along with volume and sentiment of the public news flow in explaining movements in volatility.

Both the sentiment, and rate of news flow have a pronounced impact on volatility. After controlling for expected news flow, unexpected positive shocks to the rate of news arrival, and negative shocks to the sentiment of news flow exhibit the largest impacts on volatility. Here, the impact and direction of shocks to trading activity conditional on information flow was also examined. An important finding is that the impact of all trading activity measures decrease after news is accounted for, indicating that a non-negligible component of trading is in response to public news flow. These results are found to be robust to the choice of volatility proxy. The only measure of trading activity that continues to be significant after controlling for news is trading volume. In contrast, net positions of the various types of traders are insignificant implying that the behavior of no single group of traders lead to these markets being more volatile.

References


Figure 1: Realized volatilities and information flow variables
Table 1: Descriptive Statistics and Correlation Matrix (Oil)

<table>
<thead>
<tr>
<th></th>
<th>RV</th>
<th>PK</th>
<th>R</th>
<th>S</th>
<th>N</th>
<th>VOL</th>
<th>OI</th>
<th>NPS</th>
<th>NP^H</th>
<th>NP^N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.364</td>
<td>0.333</td>
<td>0.203</td>
<td>0.074</td>
<td>2.394</td>
<td>4.407</td>
<td>11.284</td>
<td>0.024</td>
<td>-0.024</td>
<td>0.000</td>
</tr>
<tr>
<td>StDev</td>
<td>0.162</td>
<td>0.152</td>
<td>5.241</td>
<td>0.149</td>
<td>0.629</td>
<td>2.093</td>
<td>3.29</td>
<td>0.027</td>
<td>0.032</td>
<td>0.008</td>
</tr>
<tr>
<td>Min</td>
<td>0.154</td>
<td>0.119</td>
<td>-25.143</td>
<td>-0.448</td>
<td>1.132</td>
<td>0.988</td>
<td>4.542</td>
<td>-0.056</td>
<td>-0.102</td>
<td>-0.030</td>
</tr>
<tr>
<td>Max</td>
<td>1.834</td>
<td>1.200</td>
<td>21.888</td>
<td>0.353</td>
<td>3.988</td>
<td>10.47</td>
<td>16.538</td>
<td>0.088</td>
<td>0.083</td>
<td>0.027</td>
</tr>
<tr>
<td>ADF</td>
<td>-5.616*</td>
<td>-5.658*</td>
<td>-16.803*</td>
<td>-5.456*</td>
<td>-5.167*</td>
<td>-5.652*</td>
<td>-1.201</td>
<td>-3.71*</td>
<td>-3.604*</td>
<td>-4.402*</td>
</tr>
</tbody>
</table>

Correlation Matrix

<table>
<thead>
<tr>
<th></th>
<th>RV</th>
<th>PK</th>
<th>R</th>
<th>S</th>
<th>N</th>
<th>VOL</th>
<th>OI</th>
<th>NPS</th>
<th>NP^H</th>
<th>NP^N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>StDev</td>
<td>0.162</td>
<td>0.152</td>
<td>5.241</td>
<td>0.149</td>
<td>0.629</td>
<td>2.093</td>
<td>3.29</td>
<td>0.027</td>
<td>0.032</td>
<td>0.008</td>
</tr>
<tr>
<td>Min</td>
<td>0.154</td>
<td>0.119</td>
<td>-25.143</td>
<td>-0.448</td>
<td>1.132</td>
<td>0.988</td>
<td>4.542</td>
<td>-0.056</td>
<td>-0.102</td>
<td>-0.030</td>
</tr>
<tr>
<td>Max</td>
<td>1.834</td>
<td>1.200</td>
<td>21.888</td>
<td>0.353</td>
<td>3.988</td>
<td>10.47</td>
<td>16.538</td>
<td>0.088</td>
<td>0.083</td>
<td>0.027</td>
</tr>
<tr>
<td>ADF</td>
<td>-5.616*</td>
<td>-5.658*</td>
<td>-16.803*</td>
<td>-5.456*</td>
<td>-5.167*</td>
<td>-5.652*</td>
<td>-1.201</td>
<td>-3.71*</td>
<td>-3.604*</td>
<td>-4.402*</td>
</tr>
</tbody>
</table>

RV is the weekly annualized realized volatility based on 5 minute returns. PK is the weekly annualized volatility estimate based on daily Parkinson ranges. R is the weekly (Tuesday to Tuesday) return (logarithmic change of settlement prices) of the nearest futures in per cent. S, N and VOL are calculated as the average daily news sentiment, number of news and trading volume on a weekly basis (from Wednesday to Tuesday). OI is the open interest reported by CFTC relating to the outstanding futures positions as per Tuesday every week. N is in units of 100 news items. VOL and OI are in units of 100,000 futures contracts. NPS, NP^H and NP^N denote the weekly net positions of speculators, hedgers and non-reporting traders. ADF test statistics are for the hypothesis that a series contains a unit root. * indicates significance at the 1% level.

Table 2: Descriptive Statistics and Correlation Matrix (Gold)

<table>
<thead>
<tr>
<th></th>
<th>RV</th>
<th>PK</th>
<th>R</th>
<th>S</th>
<th>N</th>
<th>VOL</th>
<th>OI</th>
<th>NPS</th>
<th>NP^H</th>
<th>NP^N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.191</td>
<td>0.177</td>
<td>0.313</td>
<td>0.022</td>
<td>0.724</td>
<td>1.182</td>
<td>3.847</td>
<td>0.170</td>
<td>-0.220</td>
<td>0.050</td>
</tr>
<tr>
<td>StDev</td>
<td>0.077</td>
<td>0.088</td>
<td>2.701</td>
<td>0.156</td>
<td>0.123</td>
<td>0.679</td>
<td>1.101</td>
<td>0.046</td>
<td>0.045</td>
<td>0.018</td>
</tr>
<tr>
<td>Min</td>
<td>0.089</td>
<td>0.051</td>
<td>-11.315</td>
<td>-0.441</td>
<td>0.300</td>
<td>0.235</td>
<td>1.716</td>
<td>0.022</td>
<td>-0.321</td>
<td>0.009</td>
</tr>
<tr>
<td>Max</td>
<td>0.659</td>
<td>0.751</td>
<td>13.104</td>
<td>0.404</td>
<td>1.176</td>
<td>3.204</td>
<td>6.508</td>
<td>0.255</td>
<td>-0.075</td>
<td>0.124</td>
</tr>
</tbody>
</table>

Correlation Matrix

<table>
<thead>
<tr>
<th></th>
<th>RV</th>
<th>PK</th>
<th>R</th>
<th>S</th>
<th>N</th>
<th>VOL</th>
<th>OI</th>
<th>NPS</th>
<th>NP^H</th>
<th>NP^N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>StDev</td>
<td>0.077</td>
<td>0.088</td>
<td>2.701</td>
<td>0.156</td>
<td>0.123</td>
<td>0.679</td>
<td>1.101</td>
<td>0.046</td>
<td>0.045</td>
<td>0.018</td>
</tr>
<tr>
<td>Min</td>
<td>0.089</td>
<td>0.051</td>
<td>-11.315</td>
<td>-0.441</td>
<td>0.300</td>
<td>0.235</td>
<td>1.716</td>
<td>0.022</td>
<td>-0.321</td>
<td>0.009</td>
</tr>
<tr>
<td>Max</td>
<td>0.659</td>
<td>0.751</td>
<td>13.104</td>
<td>0.404</td>
<td>1.176</td>
<td>3.204</td>
<td>6.508</td>
<td>0.255</td>
<td>-0.075</td>
<td>0.124</td>
</tr>
</tbody>
</table>

RV is the weekly annualized realized volatility based on 5 minute returns. PK is the weekly annualized volatility estimate based on daily Parkinson ranges. R is the weekly (Tuesday to Tuesday) return (logarithmic change of settlement prices) of the nearest futures in per cent. S, N and VOL are calculated as the average daily news sentiment, number of news and trading volume on a weekly basis (from Wednesday to Tuesday). OI is the open interest reported by CFTC relating to the outstanding futures positions as per Tuesday every week. N is in units of 100 news items. VOL and OI are in units of 100,000 futures contracts. NPS, NP^H and NP^N denote the weekly net positions of speculators, hedgers and non-reporting traders. ADF test statistics are for the hypothesis that a series contains a unit root. * indicates significance at the 1% level.
<table>
<thead>
<tr>
<th></th>
<th>Gold</th>
<th>Oil</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>0.036***</td>
<td>0.032</td>
</tr>
<tr>
<td></td>
<td>(2.84)</td>
<td>(1.33)</td>
</tr>
<tr>
<td>Sum of lagged volatilities</td>
<td>0.843***</td>
<td>0.873***</td>
</tr>
<tr>
<td></td>
<td>(452.18)</td>
<td>(326.24)</td>
</tr>
<tr>
<td>Sentiment</td>
<td>-0.046***</td>
<td>-0.146***</td>
</tr>
<tr>
<td></td>
<td>(-3.91)</td>
<td>(-4.67)</td>
</tr>
<tr>
<td>Expected sentiment</td>
<td>0.038</td>
<td>-0.096*</td>
</tr>
<tr>
<td></td>
<td>(1.58)</td>
<td>(-1.94)</td>
</tr>
<tr>
<td>Unexpected sentiment</td>
<td>-0.072***</td>
<td>-0.206***</td>
</tr>
<tr>
<td></td>
<td>(-4.61)</td>
<td>(-3.90)</td>
</tr>
<tr>
<td>Unexpected sentiment</td>
<td>-0.115***</td>
<td>-0.265***</td>
</tr>
<tr>
<td></td>
<td>(-4.72)</td>
<td>(-2.60)</td>
</tr>
<tr>
<td>Number of news</td>
<td>0.089***</td>
<td>0.036***</td>
</tr>
<tr>
<td></td>
<td>(6.03)</td>
<td>(4.80)</td>
</tr>
<tr>
<td>Expected number of news</td>
<td>0.064***</td>
<td>0.026***</td>
</tr>
<tr>
<td></td>
<td>(3.2)</td>
<td>(2.14)</td>
</tr>
<tr>
<td>Unexpected number of news</td>
<td>0.108***</td>
<td>0.046***</td>
</tr>
<tr>
<td></td>
<td>(4.17)</td>
<td>(2.95)</td>
</tr>
<tr>
<td>Unexpected number of news</td>
<td>0.083***</td>
<td>0.033</td>
</tr>
<tr>
<td></td>
<td>(2.25)</td>
<td>(1.19)</td>
</tr>
<tr>
<td>Unexpected number of news</td>
<td>0.133***</td>
<td>0.060***</td>
</tr>
<tr>
<td></td>
<td>(2.26)</td>
<td>(2.03)</td>
</tr>
<tr>
<td>Durbin-Watson</td>
<td>1.927</td>
<td>2.018</td>
</tr>
<tr>
<td>Adjusted $R^2$</td>
<td>0.760</td>
<td>0.681</td>
</tr>
</tbody>
</table>

Test statistics for the sum of lagged volatilities are $F$ statistics for the hypothesis that the sum of the coefficients is zero. Test statistics for the remaining coefficients are $t$ statistics based on White heteroskedasticity consistent standard errors. ***Indicates significance at the 1% level, **denotes significance at the 5% level, and *indicates significance at the 10% level.
Table 4: Realized Volatility of Oil Futures and Shocks in Trading Activity

<table>
<thead>
<tr>
<th></th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 4</th>
<th>Model 5</th>
<th>Model 6</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Intercept</td>
<td>Intercept</td>
<td>Intercept</td>
<td>Intercept</td>
<td>Intercept</td>
<td>Intercept</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Decomposition with an ARIMA model (without news)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>0.035*</td>
<td>0.030</td>
<td>0.039*</td>
<td>0.034</td>
<td>0.034</td>
<td>0.034</td>
</tr>
<tr>
<td></td>
<td>(1.65)</td>
<td>(1.41)</td>
<td>(1.80)</td>
<td>(1.58)</td>
<td>(1.59)</td>
<td>(1.59)</td>
</tr>
<tr>
<td>Sum of lagged volatilities</td>
<td>0.901***</td>
<td>0.915***</td>
<td>0.892***</td>
<td>0.906***</td>
<td>0.905***</td>
<td>0.905***</td>
</tr>
<tr>
<td></td>
<td>(285.70)</td>
<td>(294.84)</td>
<td>(282.68)</td>
<td>(290.59)</td>
<td>(291.96)</td>
<td>(290.20)</td>
</tr>
<tr>
<td>Unexpected trading volume</td>
<td>0.026***</td>
<td>0.030***</td>
<td>0.030***</td>
<td>0.030***</td>
<td>0.030***</td>
<td>0.030***</td>
</tr>
<tr>
<td></td>
<td>(4.89)</td>
<td>(5.20)</td>
<td>(5.25)</td>
<td>(5.34)</td>
<td>(5.34)</td>
<td>(5.34)</td>
</tr>
<tr>
<td>Unexpected open interest</td>
<td>-0.047***</td>
<td>-0.056***</td>
<td>-0.056***</td>
<td>-0.058***</td>
<td>-0.058***</td>
<td>-0.058***</td>
</tr>
<tr>
<td></td>
<td>(-2.77)</td>
<td>(-3.16)</td>
<td>(-3.24)</td>
<td>(-3.57)</td>
<td>(-3.57)</td>
<td>(-3.57)</td>
</tr>
<tr>
<td>Unexpected NP speculators</td>
<td>-0.767</td>
<td>-0.767</td>
<td>-0.767</td>
<td>-0.767</td>
<td>-0.767</td>
<td>-0.767</td>
</tr>
<tr>
<td></td>
<td>(-1.38)</td>
<td>(-1.38)</td>
<td>(-1.38)</td>
<td>(-1.38)</td>
<td>(-1.38)</td>
<td>(-1.38)</td>
</tr>
<tr>
<td>Unexpected trading volume</td>
<td>0.026***</td>
<td>0.030***</td>
<td>0.030***</td>
<td>0.030***</td>
<td>0.030***</td>
<td>0.030***</td>
</tr>
<tr>
<td></td>
<td>(4.89)</td>
<td>(5.20)</td>
<td>(5.25)</td>
<td>(5.34)</td>
<td>(5.34)</td>
<td>(5.34)</td>
</tr>
<tr>
<td>Unexpected open interest</td>
<td>-0.040***</td>
<td>-0.051***</td>
<td>-0.050***</td>
<td>-0.051***</td>
<td>-0.051***</td>
<td>-0.051***</td>
</tr>
<tr>
<td></td>
<td>(-2.39)</td>
<td>(-2.94)</td>
<td>(-2.97)</td>
<td>(-3.09)</td>
<td>(-3.09)</td>
<td>(-3.09)</td>
</tr>
<tr>
<td>Unexpected NP speculators</td>
<td>-0.178</td>
<td>-0.178</td>
<td>-0.178</td>
<td>-0.178</td>
<td>-0.178</td>
<td>-0.178</td>
</tr>
<tr>
<td></td>
<td>(-0.33)</td>
<td>(-0.33)</td>
<td>(-0.33)</td>
<td>(-0.33)</td>
<td>(-0.33)</td>
<td>(-0.33)</td>
</tr>
<tr>
<td>Unexpected NP hedgers</td>
<td></td>
<td></td>
<td></td>
<td>0.817*</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(1.86)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unexpected NP small traders</td>
<td></td>
<td></td>
<td></td>
<td>-1.066</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(-0.96)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Durbin-Watson</td>
<td>2.003</td>
<td>2.105</td>
<td>1.969</td>
<td>2.091</td>
<td>2.090</td>
<td>2.093</td>
</tr>
<tr>
<td>Adjusted $R^2$</td>
<td>0.662</td>
<td>0.682</td>
<td>0.672</td>
<td>0.699</td>
<td>0.699</td>
<td>0.698</td>
</tr>
<tr>
<td>$\Delta R^2$</td>
<td>0.020</td>
<td>0.010</td>
<td>0.036</td>
<td>0.037</td>
<td>0.037</td>
<td>0.036</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Decomposition with an ARIMAX model (with news)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>0.035*</td>
<td>0.030</td>
</tr>
<tr>
<td></td>
<td>(1.65)</td>
<td>(1.41)</td>
</tr>
<tr>
<td>Sum of lagged volatilities</td>
<td>0.901***</td>
<td>0.915***</td>
</tr>
<tr>
<td></td>
<td>(285.70)</td>
<td>(291.36)</td>
</tr>
<tr>
<td>Unexpected trading volume</td>
<td>0.024***</td>
<td>0.028***</td>
</tr>
<tr>
<td></td>
<td>(4.33)</td>
<td>(4.62)</td>
</tr>
<tr>
<td>Unexpected open interest</td>
<td>-0.040***</td>
<td>-0.051***</td>
</tr>
<tr>
<td></td>
<td>(-2.39)</td>
<td>(-2.94)</td>
</tr>
<tr>
<td>Unexpected NP speculators</td>
<td>-0.178</td>
<td>-0.178</td>
</tr>
<tr>
<td></td>
<td>(-0.33)</td>
<td>(-0.33)</td>
</tr>
<tr>
<td>Unexpected NP hedgers</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unexpected NP small traders</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Durbin-Watson</td>
<td>2.003</td>
<td>2.086</td>
</tr>
<tr>
<td>Adjusted $R^2$</td>
<td>0.662</td>
<td>0.678</td>
</tr>
<tr>
<td>$\Delta R^2$</td>
<td>0.016</td>
<td>0.007</td>
</tr>
</tbody>
</table>

Trading volume and open interest are expressed in units of 100,000 contracts. Trading volume, open interest and net positions of traders are decomposed into expected and unexpected components based on an ARIMA(p,k,q) model including news sentiment and number of news items as exogenous variables. Test statistics for the sum of lagged volatilities are $F$ statistics for the hypothesis that the sum of the coefficients is zero. Test statistics for the remaining coefficients are $t$ statistics based on White heteroskedasticity consistent standard errors. $\Delta R^2$ denotes the increase in adjusted $R^2$ compared to Model 1. ***Indicates significance at the 1% level, **denotes significance at the 5% level, and *indicates significance at the 10% level.
Table 5: Realized Volatility of Gold Futures and Shocks in Trading Activity

<table>
<thead>
<tr>
<th></th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 4</th>
<th>Model 5</th>
<th>Model 6</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Decomposition with an ARIMA model (without news)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>0.016**</td>
<td>0.013*</td>
<td>0.017***</td>
<td>0.014*</td>
<td>0.015*</td>
<td>0.015*</td>
</tr>
<tr>
<td></td>
<td>(1.97)</td>
<td>(1.72)</td>
<td>(2.12)</td>
<td>(1.83)</td>
<td>(1.90)</td>
<td>(1.89)</td>
</tr>
<tr>
<td>Sum of lagged volatilities</td>
<td>0.912***</td>
<td>0.927***</td>
<td>0.907***</td>
<td>0.923***</td>
<td>0.920***</td>
<td>0.920***</td>
</tr>
<tr>
<td></td>
<td>(510.68)</td>
<td>(598.17)</td>
<td>(512.67)</td>
<td>(510.63)</td>
<td>(604.62)</td>
<td>(519.48)</td>
</tr>
<tr>
<td>Unexpected trading volume</td>
<td>0.041***</td>
<td>0.041***</td>
<td>0.041***</td>
<td>0.040***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(6.33)</td>
<td>(6.68)</td>
<td>(6.54)</td>
<td>(5.80)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unexpected open interest</td>
<td>-0.025**</td>
<td>-0.010</td>
<td>-0.011</td>
<td>-0.021**</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-2.16)</td>
<td>(-0.82)</td>
<td>(-0.96)</td>
<td>(-2.15)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unexpected NP speculators</td>
<td>-0.287**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-2.18)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unexpected NP hedgers</td>
<td></td>
<td></td>
<td>0.293**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(2.27)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unexpected NP small traders</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.527</td>
<td>(-1.23)</td>
</tr>
<tr>
<td>Durbin-Watson</td>
<td>1.989</td>
<td>2.045</td>
<td>1.958</td>
<td>2.017</td>
<td>2.016</td>
<td>2.025</td>
</tr>
<tr>
<td>Adjusted $R^2$</td>
<td>0.734</td>
<td>0.772</td>
<td>0.736</td>
<td>0.776</td>
<td>0.777</td>
<td>0.774</td>
</tr>
<tr>
<td>$\Delta R^2$</td>
<td>0.038</td>
<td>0.002</td>
<td>0.043</td>
<td>0.044</td>
<td>0.044</td>
<td>0.040</td>
</tr>
</tbody>
</table>

| **Decomposition with an ARIMAX model (with news)** |                  |                  |                  |                  |                  |                  |
| Intercept           | 0.016**          | 0.013*           | 0.016**          | 0.013*           | 0.013*           | 0.013*           |
|                     | (1.97)           | (1.68)           | (1.99)           | (1.65)           | (1.66)           | (1.69)           |
| Sum of lagged volatilities | 0.912***         | 0.927***         | 0.912***         | 0.929***         | 0.928***         | 0.927***         |
|                     | (510.68)         | (560.03)         | (518.65)         | (560.30)         | (557.98)         | (563.87)         |
| Unexpected trading volume | 0.036***         | 0.037***         | 0.036***         | 0.036***         |                  |                  |
|                     | (5.43)           | (5.75)           | (5.69)           | (5.23)           |                  |                  |
| Unexpected open interest | -0.013           | -0.004           | -0.006           | -0.010           |                  |                  |
|                     | (-1.04)          | (-0.30)          | (-0.47)          | (-0.86)          |                  |                  |
| Unexpected NP speculators | -0.221*         |                  |                  |                  |                  |                  |
|                     | (-1.76)          |                  |                  |                  |                  |                  |
| Unexpected NP hedgers |                  |                  | 0.201            |                  |                  |                  |
|                     |                  |                  | (1.64)           |                  |                  |                  |
| Unexpected NP small traders |                  |                  |                  | -0.070           |                  | (-0.17)          |
| Durbin-Watson       | 1.989            | 2.019            | 1.976            | 2.005            | 2.005            | 2.012            |
| Adjusted $R^2$      | 0.734            | 0.761            | 0.734            | 0.763            | 0.762            | 0.760            |
| $\Delta R^2$        | 0.027            | 0.000            | 0.029            | 0.029            | 0.029            | 0.027            |

Trading volume and open interest are expressed in units of 100,000 contracts. Trading volume, open interest and net positions of traders are decomposed into expected and unexpected components based on an ARIMA(p,k,q) model including news sentiment and number of news items as exogenous variables. Test statistics for the sum of lagged volatilities are $F$ statistics for the hypothesis that the sum of the coefficients is zero. Test statistics for the remaining coefficients are $t$ statistics based on White heteroskedasticity consistent standard errors. $\Delta R^2$ denotes the increase in adjusted $R^2$ compared to Model 1. ***Indicates significance at the 1% level, **denotes significance at the 5% level, and *indicates significance at the 10% level.
<table>
<thead>
<tr>
<th></th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 4</th>
<th>Model 5</th>
<th>Model 6</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Decomposition with an ARIMA model (without news)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>0.031*</td>
<td>0.026*</td>
<td>0.032*</td>
<td>0.028*</td>
<td>0.028*</td>
<td>0.028*</td>
</tr>
<tr>
<td></td>
<td>(1.90)</td>
<td>(1.71)</td>
<td>(1.95)</td>
<td>(1.82)</td>
<td>(1.84)</td>
<td>(1.87)</td>
</tr>
<tr>
<td>Sum of lagged volatilities</td>
<td>0.907***</td>
<td>0.922***</td>
<td>0.904***</td>
<td>0.915***</td>
<td>0.915***</td>
<td>0.915***</td>
</tr>
<tr>
<td></td>
<td>(348.01)</td>
<td>(399.31)</td>
<td>(346.94)</td>
<td>(383.08)</td>
<td>(391.27)</td>
<td>(411.05)</td>
</tr>
<tr>
<td>Unexpected trading volume</td>
<td>0.034***</td>
<td>0.035***</td>
<td>0.035***</td>
<td>0.036***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(6.16)</td>
<td>(6.13)</td>
<td>(6.21)</td>
<td>(6.26)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unexpected open interest</td>
<td>-0.014</td>
<td>-0.026</td>
<td>-0.025</td>
<td>-0.027*</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-0.88)</td>
<td>(-1.56)</td>
<td>(-1.58)</td>
<td>(-1.82)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unexpected NP speculators</td>
<td>-0.848</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-1.48)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unexpected NP hedgers</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.034**</td>
<td>(2.30)</td>
</tr>
<tr>
<td>Unexpected NP small traders</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-1.816</td>
<td>(-1.56)</td>
</tr>
<tr>
<td>Durbin-Watson</td>
<td>2.003</td>
<td>2.085</td>
<td>1.993</td>
<td>2.087</td>
<td>2.085</td>
<td>2.073</td>
</tr>
<tr>
<td>Adjusted $R^2$</td>
<td>0.677</td>
<td>0.716</td>
<td>0.678</td>
<td>0.721</td>
<td>0.722</td>
<td>0.720</td>
</tr>
<tr>
<td>$\Delta R^2$</td>
<td>0.038</td>
<td>0.000</td>
<td>0.043</td>
<td>0.045</td>
<td>0.045</td>
<td>0.043</td>
</tr>
<tr>
<td><strong>Decomposition with an ARIMAX model (with news)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>0.031*</td>
<td>0.026*</td>
<td>0.031*</td>
<td>0.027*</td>
<td>0.027*</td>
<td>0.027*</td>
</tr>
<tr>
<td></td>
<td>(1.90)</td>
<td>(1.71)</td>
<td>(1.91)</td>
<td>(1.73)</td>
<td>(1.72)</td>
<td>(1.74)</td>
</tr>
<tr>
<td>Sum of lagged volatilities</td>
<td>0.907***</td>
<td>0.922***</td>
<td>0.907***</td>
<td>0.920***</td>
<td>0.921***</td>
<td>0.920***</td>
</tr>
<tr>
<td></td>
<td>(348.01)</td>
<td>(394.01)</td>
<td>(347.25)</td>
<td>(386.90)</td>
<td>(389.56)</td>
<td>(400.09)</td>
</tr>
<tr>
<td>Unexpected trading volume</td>
<td>0.031***</td>
<td>0.033***</td>
<td>0.033***</td>
<td>0.033***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(5.48)</td>
<td>(5.45)</td>
<td>(5.47)</td>
<td>(5.49)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unexpected open interest</td>
<td>-0.009</td>
<td>-0.021</td>
<td>-0.020</td>
<td>-0.021</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-0.55)</td>
<td>(-1.36)</td>
<td>(-1.35)</td>
<td>(-1.43)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unexpected NP speculators</td>
<td>-0.287</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-0.52)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unexpected NP hedgers</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.549</td>
<td>(1.27)</td>
</tr>
<tr>
<td>Unexpected NP small traders</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-1.537</td>
<td>(-1.39)</td>
</tr>
<tr>
<td>Durbin-Watson</td>
<td>2.003</td>
<td>2.074</td>
<td>1.996</td>
<td>2.065</td>
<td>2.064</td>
<td>2.058</td>
</tr>
<tr>
<td>Adjusted $R^2$</td>
<td>0.677</td>
<td>0.710</td>
<td>0.677</td>
<td>0.711</td>
<td>0.712</td>
<td>0.712</td>
</tr>
<tr>
<td>$\Delta R^2$</td>
<td>0.032</td>
<td>0.000</td>
<td>0.034</td>
<td>0.034</td>
<td>0.035</td>
<td></td>
</tr>
</tbody>
</table>

Trading volume and open interest are expressed in units of 100,000 contracts. Trading volume, open interest and net positions of traders are decomposed into expected and unexpected components based on an ARIMA(p,k,q) model including news sentiment and number of news items as exogenous variables. Test statistics for the sum of lagged volatilities are $F$ statistics for the hypothesis that the sum of the coefficients is zero. Test statistics for the remaining coefficients are $t$ statistics based on White heteroskedasticity consistent standard errors. $\Delta R^2$ denotes the increase in adjusted $R^2$ compared to Model 1. ***Indicates significance at the 1% level, **denotes significance at the 5% level, and *indicates significance at the 10% level.
Table 7: Range-based Volatility of Gold Futures and Shocks in Trading Activity

<table>
<thead>
<tr>
<th></th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 4</th>
<th>Model 5</th>
<th>Model 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Decomposition with an ARIMA model (without news)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>0.019***</td>
<td>0.019***</td>
<td>0.020***</td>
<td>0.019***</td>
<td>0.019***</td>
<td>0.019***</td>
</tr>
<tr>
<td></td>
<td>(3.39)</td>
<td>(3.90)</td>
<td>(3.51)</td>
<td>(3.99)</td>
<td>(4.02)</td>
<td>(3.95)</td>
</tr>
<tr>
<td>Sum of lagged volatilities</td>
<td>0.863***</td>
<td>0.864***</td>
<td>0.859***</td>
<td>0.865***</td>
<td>0.863***</td>
<td>0.862***</td>
</tr>
<tr>
<td></td>
<td>(536.75)</td>
<td>(913.32)</td>
<td>(562.12)</td>
<td>(1023.40)</td>
<td>(981.74)</td>
<td>(896.17)</td>
</tr>
<tr>
<td>Unexpected trading volume</td>
<td>0.077***</td>
<td>0.077***</td>
<td>0.077***</td>
<td>0.077***</td>
<td>0.077***</td>
<td>0.077***</td>
</tr>
<tr>
<td></td>
<td>(8.37)</td>
<td>(8.72)</td>
<td>(8.56)</td>
<td>(8.13)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unexpected open interest</td>
<td>-0.033*</td>
<td>-0.017</td>
<td>-0.019</td>
<td>-0.024*</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-1.82)</td>
<td>(-0.90)</td>
<td>(-1.05)</td>
<td>(-1.65)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unexpected NP speculators</td>
<td>-0.227</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-1.12)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unexpected NP hedgers</td>
<td>0.077***</td>
<td>0.077***</td>
<td>0.077***</td>
<td>0.077***</td>
<td>0.077***</td>
<td>0.077***</td>
</tr>
<tr>
<td></td>
<td>(8.37)</td>
<td>(8.72)</td>
<td>(8.56)</td>
<td>(8.13)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unexpected NP small traders</td>
<td>0.077***</td>
<td>0.077***</td>
<td>0.077***</td>
<td>0.077***</td>
<td>0.077***</td>
<td>0.077***</td>
</tr>
<tr>
<td></td>
<td>(8.37)</td>
<td>(8.72)</td>
<td>(8.56)</td>
<td>(8.13)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Durbin-Watson</td>
<td>2.067</td>
<td>2.105</td>
<td>2.044</td>
<td>2.102</td>
<td>2.102</td>
<td>2.098</td>
</tr>
<tr>
<td>Adjusted $R^2$</td>
<td>0.638</td>
<td>0.717</td>
<td>0.641</td>
<td>0.719</td>
<td>0.719</td>
<td>0.718</td>
</tr>
<tr>
<td>$\Delta R^2$</td>
<td>0.079</td>
<td>0.002</td>
<td>0.081</td>
<td>0.081</td>
<td>0.081</td>
<td>0.081</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 4</th>
<th>Model 5</th>
<th>Model 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Decomposition with an ARIMAX model (with news)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>0.018***</td>
<td>0.018***</td>
<td>0.019***</td>
<td>0.018***</td>
<td>0.018***</td>
<td>0.018***</td>
</tr>
<tr>
<td></td>
<td>(3.68)</td>
<td>(3.33)</td>
<td>(3.59)</td>
<td>(3.61)</td>
<td>(3.62)</td>
<td></td>
</tr>
<tr>
<td>Sum of lagged volatilities</td>
<td>0.868***</td>
<td>0.864***</td>
<td>0.872***</td>
<td>0.871***</td>
<td>0.869***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(843.91)</td>
<td>(549.38)</td>
<td>(911.92)</td>
<td>(901.57)</td>
<td>(849.73)</td>
<td></td>
</tr>
<tr>
<td>Unexpected trading volume</td>
<td>0.073***</td>
<td>0.073***</td>
<td>0.073***</td>
<td>0.073***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(7.48)</td>
<td>(7.83)</td>
<td>(7.73)</td>
<td>(7.42)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unexpected open interest</td>
<td>-0.021</td>
<td>-0.011</td>
<td>-0.014</td>
<td>-0.013</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-1.03)</td>
<td>(-0.57)</td>
<td>(-0.71)</td>
<td>(-0.77)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unexpected NP speculators</td>
<td>-0.170</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-0.93)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unexpected NP hedgers</td>
<td>0.123</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.70)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unexpected NP small traders</td>
<td>0.348</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.61)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Durbin-Watson</td>
<td>2.067</td>
<td>2.102</td>
<td>2.055</td>
<td>2.099</td>
<td>2.098</td>
<td>2.096</td>
</tr>
<tr>
<td>Adjusted $R^2$</td>
<td>0.638</td>
<td>0.705</td>
<td>0.639</td>
<td>0.705</td>
<td>0.705</td>
<td>0.705</td>
</tr>
<tr>
<td>$\Delta R^2$</td>
<td>0.067</td>
<td>0.000</td>
<td>0.067</td>
<td>0.067</td>
<td>0.067</td>
<td>0.066</td>
</tr>
</tbody>
</table>

Trading volume and open interest are expressed in units of 100,000 contracts. Trading volume, open interest and net positions of traders are decomposed into expected and unexpected components based on an ARIMA(p,k,q) model including news sentiment and number of news items as exogenous variables. Test statistics for the sum of lagged volatilities are $F$ statistics for the hypothesis that the sum of the coefficients is zero. Test statistics for the remaining coefficients are $t$ statistics based on White heteroskedasticity consistent standard errors. $\Delta R^2$ denotes the increase in adjusted $R^2$ compared to Model 1. ***Indicates significance at the 1% level, **denotes significance at the 5% level, and *indicates significance at the 10% level.