WORKING PAPER SERIES

WORKING PAPER
N. 145
JULY 2020

Media Attention vs. Sentiment as Drivers of Conditional Volatility Predictions: An Application to Brexit

By Massimo Guidolin and Manuela Pedio

Università Bocconi
BAFFI CAREFIN
Centre for Applied Research on International Markets, Banking, Finance and Regulation

Media Attention vs. Sentiment as Drivers of Conditional Volatility Predictions: An Application to Brexit

Massimo Guidolin[♣]

Bocconi University, Baffi CAREFIN

Manuela Pedio

Bocconi University, Baffi CAREFIN

July 2020

Abstract

Using data on international, on-line media coverage and tone of the Brexit referendum, we test whether it is media coverage or tone to provide the largest forecasting performance improvements in the prediction of the conditional variance of weekly FTSE 100 stock returns. We find that versions of standard symmetric and asymmetric Generalized Autoregressive Conditional Heteroskedasticity (GARCH) models augmented to include media coverage and especially media tone scores outperform traditional GARCH models both in- and out-of-sample.

Keywords: Attention, Sentiment, Text Mining, Forecasting, Conditional Variance, GARCH model, Brexit.

JEL codes: C53, C58, G17.

1. Introduction

It is now well understood that media coverage (taken as a proxy for the *attention* that market operators are devoting to a phenomenon, asset, or event) and tone (a proxy for market *sentiment*) are important drivers of trading volumes, stock returns, and to some extent, also risk (see Brown and Cliff, 2005; Da, Engelberg, and Gao, 2011; Fang and Peress, 2009; Tetlock, 2007). While the evidence on the predictive power of media coverage and sentiment for stock returns and trading volume is now rather pervasive, the

[♣] Corresponding author. Address: Bocconi University, Finance dept., via Roetgen 1, 20136 Milan, Italy. E-mail: massimo.guidolin@unibocconi.it. Tel.: +39-02-5836.5885.

available empirical results concerning the connection between media attention and tone and standard measures of risk, such as conditional variance, were initially tenuous at best. For instance, Mitchell and Mulherin (1994) and Berry and Howe (1994) tried to relate aggregate stock market volume and volatility to broad measures of news about firms and the economy (the number of Dow Jones Newswire and Wall Street Journal stories per day and the frequency of hourly news items from Reuters' News Service, respectively). Both studies reported weak correlations of less than 0.12 between market volatility and the news items. However, Antweiler and Frank (2004) studied the frequency and tone of stock message-board posts on Yahoo! Finance and Raging Bull for 45 large US stocks in the year 2000. Their main finding is that message-board posting frequency positively predicts stock return volatility and trading volume, even controlling for the frequency of WSJ stories.

In this paper, we perform analyses that to the best of our knowledge are new to test whether on-line media attention and sentiment have any predictive power for the risk, measured as the conditional variance of returns, of the FTSE 100 stock index. We base our investigation on the frequency and tone extracted from more than half a million on-line news items concerning the Brexit referendum vote (that took place in June 2016) over a daily 2015-2018 sample, available from Global Database of Events, Language, and Tone (GDELT). Importantly, we conduct our analysis within a classical Generalized Autoregressive Conditional Heteroskedasticity (GARCH) framework, i.e., we seek to measure whether and how the in- and out-sample predictive performance of off-the shelf GARCH can be improved by augmenting the model to include exogenous scores that reflect media coverage and tone.

Importantly, there is some theoretical backing four our empirical tests. For instance, Andrei and Hasler (2014) have presented an asset-pricing model that provides microeconomic foundations behind the link between asset prices, as well as volatility, and investors' attention. In their model, the level of attention varies with economic conditions, and the model shows that higher attention and higher uncertainty are associated with higher stock return variance. Andrei and Hasler propose that when investors pay more attention to news, information gets incorporated into prices faster, which results in greater market volatility. This link between attention and volatility is supported by the empirical literature (see, e.g. Da, Engelberg, and Gao, 2011; Dimpfl and Jank, 2016;

Goddard, Kita, and Wang, 2015; Vlastakis and Markellos, 2012). Paradoxically, we know much less about the link between news sentiment and stock market risk, even though Da, Engelberg, and Gao (2015) construct a new measure of investor sentiment, a Financial and Economic Attitudes Revealed by Search (FEARS) index, by aggregating the volume of Google Trends queries related to concerns such as recession, unemployment and bankruptcy. They report that FEARS also predicts temporary increases in volatility (measured as either realized volatility on the S&P500 exchange traded funds return or the market volatility index, the VIX).

In the perspective of showing the importance of attention and sentiment in forecasting the volatility of asset returns, the Brexit vote represents a highly relevant case for several reasons. Indeed, Brexit has undoubtfully raised concerns among investors: Kierzenkowski, Pain, Rusticelli, and Zwart (2016) and Dhingra, Ottaviano, Sampson, and Van Reenen (2016), among the others, estimate a perspective, strong negative economic impact for the United Kingdom in terms of reduction of their gross domestic product and of foreign investment inflows. Therefore, in this paper we argue in favor of identifying the uncertainty regarding the outcome of the referendum first, and the doubts concerning the exit strategy later, as clear factors of long-term tension, both in an economic and social perspectives. The impact of Brexit on the volatility of the stock market is confirmed by Raddant and Karimi (2016), who have examined the effects of the results of the Brexit vote on the correlation and volatility of the European financial markets, represented by five major equity indices. They show that the outcome of the Brexit vote led to a general increase in volatility, especially in the United Kingdom and in Italy, although the effect only lasted for three weeks. Belke, Dubova, and Osowski (2018) have proposed a political uncertainty index directly linked to the political dispute process around Brexit in the year preceding the consultation and used a Granger-causality test to show that part of the volatility was transmitted from the political uncertainty regarding Brexit to the financial and currency markets.

Our key result is that while contemporaneous findings that media attention increases the conditional variance of asset returns are confirmed, it is mostly the (absolute value of the) tone of the on-line media flow that generates risk.[1] Therefore, it is not just the presence of media chatter and coverage that generates risk, but also (or especially) the fact the news come to reflect strong views, with a specific tone that seems to affect stock prices more, thus offering precise forecasting power. In particular, the tone of media coverage of the Brexit is a significant additional variable when used to augment a standard GARCH model. Conversely, the level of media attention to the topic – measured by the growth rate of the articles where the word Brexit appears – is also significant, but the predictive performance is better than the one of the benchmark only when the variance of returns is high, i.e., in a specific market regime.

The use of measures of news coverage (and their tone) to predict market returns and volatility has been largely debated in the literature although the results are yet inconclusive. In his seminal paper, Tetlock (2007) first constructed an index of pessimism characterizing traders by analyzing the semantics used in the short comments contained in the popular column "Abreast of the Market" in the Wall Street Journal. He finds that high media pessimism predicts downward pressure on market prices followed by a reversion to fundamentals. Similarly, Chen, De, Hu, and Hwang (2014) have investigated the extent to which opinions transmitted through social media platforms may predict future stock market returns and report that the views expressed in the articles are able to forecast stock returns and earnings surprises. Schmidt (2013) uses Google searches for international sporting events to test a theory in which distracted investors prioritize market news over firm-specific news. He shows that a standardized increase in investor attention to sports—implying inattention to stocks—reduces market volatility by 8% and trading activity by 4%.

Guidolin, Orlov, and Pedio (2017) use the volume of internet searches (or, alternatively, its growth rate) retrieved from Google Trends as a proxy for investors' interest in a specific asset to forecast the returns of the UK, French, German, and US stock and bond markets and find that such heuristic models perform better than the average of all the alternative predictive models. Blasco, Corredor and Santamaria (2005) study what kind

[1] For instance, Wohlfarth (2018) uses EGARCH models extended to include daily Google Trends-based monetary policy attention indices and reports that policy attention has a significant impact on the variance processes of bonds whilst mean processes are just an affine function of implied volatility as captured by VIX. Shen, Urquhart, and Wang (2019) have examined the links between investor attention, measured by Twitter data, and Bitcoin returns, trading volume and realized volatility. They find that the number of tweets Granger causes both next day trading volume and realized volatility.

of information affects close-to-open and open-to-close returns, the volatilities, and the volume of actively traded securities in the Spanish market. Their study is close to ours as they also propose several specifications of the GARCH model that accounts for the asymmetric effects of the news on the stock market (conditional) volatility. However, they are not interested in the predictive accuracy of their models but only in the explanatory power of the news-related variables.

Our study contributes to this literature in several ways. First, to the best of our knowledge, we are the first to employ the extensive information coming from the Global Database of Events, Language, and Tone (GDELT) to forecast stock return volatility. This is an enormous data repository collecting and analyzing the textual content of news coming from hundreds of thousands different sources in 100 different languages. The unique features of this database allow us to gain an unprecedented coverage of the topic and to quantify both the growth of the attention devoted to Brexit and the tone attributed to the event by the commentators. Second, we propose and evaluate several extensions of the traditional GARCH model that allows the news-related variable to enter in the prediction of future conditional volatility in asymmetric ways.

Ferguson, Philip, Lam, and Guo (2015) have examined whether tone (positive and negative) and volume of firm-specific *press news* content provide valuable information about future stock returns.[2] Even though they fail to refer specifically to the Brexit referendum, their results indicate that both tone and volume of news media content significantly predict next period abnormal returns, with the impact of volume more pronounced than tone. The predictive power of tone is found to be stronger among lower visibility firms. A simple news-based trading strategy produces statistically significant risk-adjusted returns of 14.2 to 19 basis points in the period 2003--2010. At the aggregate level, price pressure induced by semantics in news stories is corrected only in part by subsequent reversals. Of course, by focussing on a recent sample, an iconic, path-breaking event (Brexit) and especially on all on-line news captured by the Global Database of Events, Language, and Tone (GDELT), we provide fresher, in our view, more cogent empirical insights on whether tone or attention ought to impact conditional risk

[2] They source the news articles from national newspapers that are globally recognised, namely, The Financial Times (FT), the Times, the Guardian and Mirror and construct the indicators from Loughran and McDonald's (2011) financial-news-specific word lists.

measures.

The rest of the paper is structured as follows. Section 2 describes the data used in the analysis and the alternative models that we propose to forecast conditional volatility. Section 3 discusses the empirical results. Section 4 concludes.

2. Data

The main source of data for our analysis is the *Global Database of Events, Language, and Tone* (henceforth, GDELT).[3] As per the definition of the maintainers, the GDELT Project is "a real time open data global graph over human society as seen through the eyes of the world's news media". More specifically, GDELT collects news from print, broadcast, and online media in more than 100 languages and from across all countries in the world and analyze them using a set of natural language processing algorithms to detect network connections and to decode reactions and emotions connected to a given event. Although the GDELT Project offers a wide range of analysis tools, to construct the series employed in this paper we have relied on the *Global Knowledge Graph* tool (henceforth, GKG). The GKG is a large collection of press articles organized around a "*nameset*", which represents a record in the database. A *nameset* is a set of names (of people, such as politicians, or firms and organizations, such as Coca-Cola or the European Union), places (like the United Kingdom or London) and themes (from a list compiled by the GDELT maintenance team) that tend to appear together in the articles. Each record contains, in addition to the number of articles that mention the *nameset* (i.e., that contain all the words composing the *nameset* at least once), a range of sentiment scores that summarizes the average tone of the articles mentioning the *nameset*. A more detailed description of the database together with a small subsample of the data can be found in Appendix A.

The GKG tool allows a researcher to extract the records referring to articles that contain one (or more) keywords. In our case, because we are interested in the media coverage of the Brexit referendum vote, we perform a search based on the word "Brexit" over a

[3] More information about the GDELT Project can be found at https://www.gdeltproject.org.

sample period spanning from October 1, 2015 to October 31, 2018.[4] The results of the extraction from the GDELT database include more than half a million articles written in the period under consideration, with peaks of more than 10,000 articles in the days close to the referendum, as shown in Figure 1.

As the goal of our analysis is to capture both the attention devoted to Brexit by the press and the "blogosphere" and the sentiment surrounding the topic, we construct two series, one measuring the growth rate of the articles and the other capturing the average tone of the coverage.[5] More precisely, the growth rate of the articles (that we shall denote as *Art Growth*) is computed as

$$Art\ growth_t = \frac{\omega_t - \omega_{t-1}}{\omega_{t-1}}\ , \tag{1}$$

where $\omega_t = \sum_{i=1}^{I} \omega_{i,t}$, i indicates a given *nameset*, I is the total number of *namesets* at date t (corresponding to the total number of records for that day), and $\omega_{i,t}$ is the number of articles mentioning *nameset i* (corresponding to the field "NUMARTS" in the GDELT database, as detailed in Appendix A). Figure 1 shows the growth rate of the article over time (note that the data are expressed in percentages; therefore 1 should be read as 100% and so on). As expected, it peaks around the date of the Referendum in June 2016. However, the chart also unveils a number of other spikes, especially during 2017; for instance, in March 2017, the British government formally announced the decision to leave the European Union; in June 2017 new elections were held in Great Britain.

The *tone* variable is instead constructed as the weighted average of the tone score of each *nameset i* published at time t, assuming as weights the total number of articles citing *nameset i*:

---

[4] Although according to the Oxford English Dictionary (2017) the word "Brexit" first appeared in May 2012, we start our analysis from October 1, 2015 because this date corresponds to the implementation of the GKG tool. This is reasonable because a closer media coverage of the topic has begun well in advance of the Referendum (June 2016). In addition, we exclusively use the word "Brexit" as a keyword because we want to avoid capturing unrelated articles (for instance, articles mentioning the "European Union" but without any connection with the main object).

[5] Da, Engelberg, and Gao (2011) proposed the identification between investors' attention and growth of the on-line media traffic. They assumed that the frequency of Google searches (Search Volume Index or SVI) for a stock's ticker is a measure of investor attention to the stock. Using a sample of US stocks from 2004 to 2008, they show that SVI positively predicts three empirical proxies for attention: news stories, trading volume, and the absolute value of stock returns.

$$tone_t = \frac{\sum_{i=1}^{I} tone_{i,t} * \omega_{i,t}}{\sum_{i=1}^{I} \omega_{i,t}} \ , \tag{2}$$

where $tone_{i,t}$ represents the tone score of the *nameset* i, and $\omega_{i,t}$ as the same meaning as above. [6] In order to align the scale of the variable with the target variable, i.e., the variance of price returns (to facilitate the functioning of the numerical maximization algorithms used in the estimation of the coefficients) both series have been scaled multiplying them by a factor of 100. Figure 3 shows the evolution of the tone variable over the sample. As expected, the tone is on average negative and stable over time with negative peaks around the date of the Referendum and of the British elections in 2017; interestingly, there is also a negative spike in January 2018, when the European Council firstly adopted and published the negotiating directives.

As our application aims at forecasting the conditional variance (volatility) of the FTSE100 Index, which is a latent variable, to evaluate our predictions, we need a proxy for its observed value. To this purpose, we employ the daily series of the realized variances based on both the open-to-close (OtC) and close-to-open (CtO) returns of the FTSE 100 index measured at a frequency of 10 minutes, without sub-sampling. The data are collected from the "Realized Library" database, developed by the Oxford-Man Institute of Quantitative Finance of the University of Oxford (UK).[7] The series of the daily CtO and OtC returns are retrieved from the Oxford-Man Institute dataset as well.

The realized volatilities of the OtC and the CtO FTSE100 returns are plotted in Figure 4 and Figure 5, respectively. The two series tend to show approximately the same pattern, although CtO returns are obviously less volatile than OtC ones as overnight trading is by construction much less intense and the news flows muted. In Figure 4, we can observe

[6] More precisely, $tone_{i,t}$ is a number between -100 (extremely negative) and +100 (extremely positive) computed as positive minus negative score. The positive score is a number between 0 and 100 representing the proportion of words with a positive emotional connotation over the total. The negative score is a number between 0 and 100 representing the proportion of words with a negative emotional connotation over the total.

[7] The Realized Library includes measurements of the daily realized variances of 31 different stock indexes, obtained according to the most widespread high-frequency econometric modelling methods proposed in the literature. The input for the calculation of the realized variances is obtained from the Thomson Reuters' database that collects historical data at an intraday frequency, namely DataScope Tick History. More detailed information about estimators and their calculation methodology can be found at https://realized.oxford-man.ox.ac.uk.

two notable spikes of the realized volatility: one in correspondence of the Brexit Referendum and one in January 2018.

Table 1 shows the summary statistics of the data and Table 2 reports the pair-wise correlation between the variables. As already mentioned, the average tone is negative and equal to -1.46 (which means that 1.46% of the total words in the text have a negative emotional connotations) with peaks below -11. The growth in the number of articles was on average 11.35% but it is extremely volatile with peaks of 600% and more. As we may expect, the tone and the article growth series are negatively correlated (i.e., an increase in the press coverage tends to carry a negative tone, as humans pay more attention to negative news and on-line sources cater to that need). In addition, the growth in the number of articles shows a large, positive correlation with the contemporaneous realized volatility; the tone displays instead a negative (but small) correlation (i.e., a more pessimistic press coverage correlates with an increase in the realized volatility). Of course, structural time series modelling, in the form of GARCH, is needed to exactly quantify the structure of any dynamic linkages, which is what we do next.

3. Methodology

3.1 News augmented GARCH models

The goal of our paper is to assess whether extending the standard GARCH model developed by Bollerslev (1986) to capture the extent (attention) and the tone (sentiment) of the media coverage of the Brexit helps predicting conditional variance of the OtC (and, as a robustness check, the CtO) daily returns of the FTSE 100. The first specification that we propose is a GARCHX obtained by simply adding each of the two news-related series as an exogenous variable (similar to Engle, 2002) to a standard GARCH:

$$R_{t+1} = \sigma_{t+1|t} z_{t+1} \qquad z_{t+1} \sim IID\ N(0,1)\,. \tag{3}$$

$$\sigma^2_{t+1|t} = \omega + \alpha R_t^2 + \beta \sigma^2_{t|t-1} + \gamma x^2_{i,t} \quad i = 1,2 \tag{4}$$

where R_t is the (daily) log return defined as $R_t \equiv \ln\left(\frac{P_t}{P_{t-1}}\right)$, P_t is the value of the FTSE 100 index at time t, z_t is a normally distributed random shock, and $\sigma^2_{t|t-1}$ is the conditional variance at time t given the information set at time t – 1, $x_{i=1,t} = tone_t$, while $x_{i=2,t} =$

Art_growth_t; ω, α, β and γ are the estimable parameters.[8]

While the news-related variables are squared in the model presented in (3)-(4), we also consider the possibility that their sign could contain important information. For instance, Campbell and Hentschel (1992) report that negative news has a larger impact on volatility than positive ones. To accommodate this asymmetric effect, we propose the following specification

$$\sigma^2_{t+1|t} = \omega + \alpha R_t^2 + \beta\sigma^2_{t|t-1} + \gamma(1 + d_1\theta)tone_t^2, \quad (5)$$

where d_1 is a dummy variable that is activated when $tone_t < 0$, $\theta\ is\ restricted\ to\ positive$, and the returns are modeled as in (3). In this way, given two news with the same impact in terms of absolute value, the most pessimistic news has a greater impact on the return variance than the most optimistic one.[9]

For what concerns the growth in the number of articles, previous literature (see e.g., Guidolin, Orlov, and Pedio, 2018) shows that an increase in the coverage of a topic will have an impact on conditional variance, while a decrease by the same amount will not. Therefore, we specify a GARCH-X model including the growth of the news items:

$$\sigma^2_{t+1|t} = \omega + \alpha R_t^2 + \beta\sigma^2_{t|t-1} + \gamma d_2 Art\_growth_t^2\ , \quad (6)$$

where d_2 is a dummy variable that is activated when $Art\_growth_t > 0$, γ is restricted to be positive, and the returns are modeled as in (3). The models in (5) and (6) are denoted as *GARCH Asymmetric News Dummy* (GARCHAND).

Finally, we conjecture that the news-related variables may only have impact on stock variance when its level is already high. Therefore, we also propose a GARCH where the exogenous variable (either the growth in the number of news items or the tone of the media) enters the model when the predicted variance for the previous period exceeded a threshold, κ. We shall call this model GARCH *News Dummy* (GARCHND):

$$\sigma^2_{t+1|t} = \omega + \alpha R_t^2 + \beta\sigma^2_{t|t-1} + \gamma d_3 x^2_{i,t} \qquad i = 1,2 \quad (7)$$

---

[8] In our analysis, we restrict ourselves to GARCH(1,1) and its extensions; as often discussed in the literature (see, e.g., Hansen and Lunde, 2005) adding lags of past squared errors and/or the predicted variance tends not improve much the forecasting performance of the model.

[9] This specification is similar to the GJR-GARCH proposed by Glosten, Jagannathan, and Runkle (1993), with the difference that the asymmetric effect concerns the exogenous variable and not the past errors.

where $x_{i=1,t} = tone_t$ and $x_{i=2,t} = Art\_growth_t$, the returns are modeled as in (3) and d_3 is a dummy variable that is activated when $\sigma^2_{t|t-1} \geq \kappa$; we experiment with two levels of κ, corresponding to a volatility of 10% and 20% (and of 1% and 2% for the volatility of the CtO returns, as a robustness check).

3.2 The estimation method

The estimation of a the GARCH model and of the extensions that represent the core of our analysis is performed using the *Maximum Likelihood Estimation* method (MLE). More specifically, we assume that shocks in (3) are IID and follows a normal distribution and maximize the (log of) the likelihood:

$$\ln L = \sum_{t=1}^{T} \ln(l_t) = \sum_{t=1}^{T} \left[-\frac{1}{2} \ln(2\pi\sigma_t^2) - \frac{R_t^2}{2\sigma_t^2} \right] \tag{8}$$

using numerical maximization algorithms that bypass the need to have analytical forms for the calculation of the derivatives of the log-likelihood function. Notably, we have achieved a success rate close to 100% in maximizing the likelihood function by implementing a "grid" of initial attempts to estimate coefficients.[10]

Once the model coefficients are estimated, in order to test their statistical significance we calculate the standard errors of the estimator $\widehat{\boldsymbol{\theta}}$, where $\boldsymbol{\theta}$ is the vector of the coefficients in the model such that the log-likelihood function, treated as a parametric function based on $\boldsymbol{\theta}$, reaches its maximum. The asymptotic distribution of the (quasi) maximum likelihood estimator is known to be:

$$\widehat{\boldsymbol{\theta}} \overset{a}{\to} N(\boldsymbol{\theta}_0, I(\boldsymbol{\theta}_0)^{-1}), \tag{9}$$

where $\boldsymbol{\theta}_0$ and $I(\boldsymbol{\theta}_0)^{-1}$ are, respectively, the (asymptotic) mean and the covariance matrix of the distribution of the estimator. More specifically, the covariance matrix of the QMLE estimator can be computer by substituting $\boldsymbol{\theta}_0$ with $\widehat{\boldsymbol{\theta}}$ and by inverting the estimated Fisher information matrix $I(\widehat{\boldsymbol{\theta}})$:

[10] The exploitation of a grid of so-called initial guesses, i.e., a range of initial values of the coefficients from which the numerical minimization algorithms start the estimation, allows to overcome the stalemate due to unsuccessful optimizations simply by trying a different combination of the same. In our recursive exercises, in the rare cases in which the grid does not solve the problem by satisfying necessary and sufficient first- and second-order conditions, the algorithm maintains the estimates of the previous period.

$$var(\widehat{\boldsymbol{\theta}}) = -I^{-1}(\widehat{\boldsymbol{\theta}}) = -\left(E\left[\frac{\partial^2 \ln \mathrm{L}(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'}\right]_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}}\right)^{-1}. \tag{10}$$

Because in most practical applications the expected value of the second derivative of the maximized log-likelihood function is not easy to compute, we follow Bollerslev and Wooldridge (1992) and calculate the *Outer Product of Gradients* (OPG) matrix, defined as:

$$J(\boldsymbol{\theta}) = \sum_{t=1}^{T} \frac{\partial \ln l_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \frac{\partial \ln l_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'}, \tag{11}$$

whose variance-covariance matrix is:

$$var(\widehat{\boldsymbol{\theta}}) = J^{-1}(\widehat{\boldsymbol{\theta}}). \tag{12}$$

The OPG matrix ensures that J is positively defined and therefore it is possible to easily obtain the standard error of the MLE estimator for any coefficient i in the model.[11]

3.3 Forecasting performance measures

Because the goal of our analysis is to assess the predictive ability of the news-augmented models, we propose a range of performance measures that are standard in the forecasting literature. As already emphasized, the target variable remains unobservable even *ex-post*; therefore, the evaluation should be based on an observable variable that may be a reliable proxy for the true and unobserved value of the variance (in our case, realized variance). We call direct methods those statistics in which the evaluation requires a direct comparison between the forecast of the target variable with the *ex-post* measure of the proxy chosen. On the contrary, we define mixed methods those that involve the comparison between two alternative predictive models, for instance an augmented GARCH and a standard GARCH benchmark. These methods are mixed because they are preliminarily based on a direct comparison between the forecasts from each model and the proxy of the target variable, then on a comparison between the respective forecast performances. More specifically, we compute the *Root Mean Squared Error* (RMSE), the *Mean Absolute Error* (MAE) and the *Mean Absolute Percentage Error* (MAPE) (direct methods) and the *Relative MAPE* (RMAPE), the Theil's *U*, and the *Diebold-Mariano* (DM) test (mixed methods).

[11] All models presented in section 3 have been written and estimated in Python language through the Integrated Development Environment (IDE) "Spider".

The RMSE is computed as

$$RMSE = \frac{1}{T}\sqrt{\sum_{t=1}^{T}\left(RV_{t+1} - \sigma_{t+1|t}^{2}\right)^{2}}\,, \tag{13}$$

where RV_{t+1}, i.e. the realized variance, is the value of the proxy chosen for the variance at time $t+1$ and $\sigma_{t+1|t}^{2}$ is the conditional variance forecasted by the model for the same day. The MAE is the average error in absolute terms and it is computed as:

$$MAE = \frac{1}{T}\sum_{t=1}^{T}\left|RV_{t+1} - \sigma_{t+1|t}^{2}\right|. \tag{14}$$

The MAPE represents the average of the absolute error in percentage and is computed as:

$$MAPE = \frac{1}{T}\sum_{t=1}^{T}\frac{\left|RV_{t+1} - \sigma_{t+1|t}^{2}\right|}{RV_{t+1}}. \tag{15}$$

In contrast with the two previous measures, MAPE does not preserve the scale of the predicted variable but ranges between zero and one. In addition to the previous statistics, we also consider a measure that compares the MAPE of each augmented model with that of the standard GARCH benchmark. This index, which ranges between zero and one, is a relative measure of forecast accuracy that compares two competing models; a value lower than one signals that the augmented model outperforms the benchmark (i.e., it has a lower mean absolute percentage error). The relative MAPE (RMAPE) is computed as:

$$RMAPE_{model} = \frac{MAPE_{model}}{MAPE_{benchmark}}\,. \tag{16}$$

The statistics U, originally introduced by Theil (1966), is defined as:

$$Theil\ U = \sqrt{\frac{\frac{1}{T}\sum_{t=1}^{T}\left(\frac{RV_{t+1} - \sigma_{t+1|t}^{2}}{RV_{t}}\right)^{2}}{\frac{1}{T}\sum_{t=1}^{T}\left(\frac{RV_{t+1} - f_{bt+1}}{RV_{t}}\right)^{2}}}\,, \tag{17}$$

where f_{bt+1} is the conditional variance predicted by a benchmark model. Theil's U is a measure of relative forecast accuracy vs. a benchmark. In this case, we choose to compare the forecasts from each predictive model vs. a *naïve* benchmark that set the forecast of future conditional variance equal to the realized variance of the previous period, i.e., $f_{bt} = RV_{t-1}$. A value of Theil's U smaller than one indicates that the predictive model outperforms the benchmark and viceversa.

Finally, Diebold and Mariano's (DM, 1995) test evaluates the null hypothesis of equal accuracy of two competing models using the t-ratio statistic computed as:

$$DM = \frac{\bar{d}}{\sqrt{\frac{\widehat{LRV}_{\bar{d}}}{T}}} \quad , \tag{18}$$

where $\bar{d} = \sum_{t=1}^{T} d_t/T$, $d_t = g(e_{i,t}) - g(e_{j,t})$, and $LRV_{\bar{d}} = \gamma_0 + 2\sum_{j=1}^{T-1} \gamma_j$ with $\gamma_j = cov(d_t, d_{t-j})$. In our application, we have chosen two different specifications of the loss function $g(\cdot)$, namely:

$$g(e_{i,t+1}) = \left(RV_{i,t+1} - \sigma_{i,t+1|t}^2\right)^2 , \tag{19}$$

and $$g(e_{i,t}) = \left|RV_{i,t+1} - \sigma_{i,t+1|t}^2\right| . \tag{20}$$

The null hypothesis of the DM test is

$$H(0): E(\bar{d}) = 0 \ \Rightarrow \ H(0): DM = 0 \tag{21}$$

Diebold and Mariano show that, under the null hypothesis of equal forecasting accuracy H_0, $DM \sim N(0,1)$. Therefore, we reject $H(0): E(\bar{d}) = 0$, i.e., the hypothesis of equal predictive accuracy, if:

$$\widehat{DM} < -z_\alpha , \tag{22}$$

where z_α is the quantile of the standard Normal distribution such that $P(DM < -z_\alpha) = \alpha$ with α being the desired size of the test.

The choice of using two different functions forms for $g(\cdot)$,the square of the forecast error and its absolute value, respectively, is dictated by two reasons. First, entertaining two loss functions enriches the information coming from the test, allowing us to compare the results of the statistics under different assumptions. In addition, it is useful to evaluate the performance of a model considering the two alternatives: a loss function that is "punitive" (the squared one, which gives a higher weight to more significant errors in absolute value and shrinks the weight of small errors) and one (the absolute value) which maintains the magnitude of errors without any asymmetric transformation based on whether an errors falls above or below one.

4. Empirical Results

4.1 In-sample forecasting results

Table 3 shows the in-sample estimates and the measures of in-sample fit for the models discussed in Section 3. Notably, the models including the growth in the number of articles either in a symmetric or an asymmetric fashion fail to utperform a standard GARCH(1,1) model, in-sample. This is also confirmed by the fact that the γ coefficient associated to the article growth variable is never statistically significant. Conversely, there are two models that include the tone variable which outperform a standard GARCH model both in terms of RMSE and of MAE: a GARCHX model where the tone is considered irrespective of its sign and a GARCHAND model in which a negative score for the tone receives more weight than a positive one. In the first case, the γ coefficient associated with tone is statistically significant and equal to 1.19; on the contrary, the γ and θ coefficients of the GARCHAND model are not statistically significant; this is likely due to the large estimation error.

From these results it appears that "sentiment", as captured by the tone of media coverage of Brexit, helps to predict the conditional variance of (OtC) stock returns over the sample period, while "attention" as measured by the growth in the number of articles does not. However, we conjecture that result may derive from the fact that while an increase in the attention is likely to be immediately compounded by the market into prices, a change in the tone (for instance, an increase in the pessimism surrounding the outcomes of Brexit) may require a longer time to be detected and processed.

To assess whether our intuition has merit, we re-estimate the models in (3)-(7) by using the news-related variable at time $t + 1$ to predict the variance at $t + 1$ (similar to Blasco et al., 2005). The estimates for these alternative models and the measures of their in-sample fit are reported in Table 4. Notably, the results show that both the growth in the number of articles and the tone variable help to explain the forecast of the conditional variance at time $t + 1$ irrespective to whether their sign is accounted for or not. As far as the growth in the number of news is concerned, the best model seems to be a GARCHND with a threshold of 10%, which outperforms a standard GARCH(1,1) both in terms of RMSE and of MAE. Similarly, also a GARCHND model that includes the tone variable when the volatility in the previous period was higher than 10% outperforms all the other models including the tone-augmented one, as well as the standard GARCH(1,1). It is worthwhile to note that when the news-related variables are measured at time $t + 1$ the models based on the growth in the number of media releases also outperform the ones based on the tone variable. This further supports the intuition that – while an increase in attention is immediately processed by market participants – a change in sentiment takes more time to be processed but can be truly exploited to make predictions about future conditional variance in real time.

4.2 Out-of-Sample Forecasting Results

Because a trader (or risk-manager) would be interested to exploit news-based forecasts only out-of-sample (OOS), in Table 5, we report the OOS forecasting performance of the four best augmented models.[12] In particular, we perform a *genuine* OOS exercise: we use the data from October 1, 2015 to June 30, 2018 in order to estimate the models; therefore, we exploit the estimated models to forecast the conditional variance for the period spanning from July 1, 2018 to October 30, 2018. This means that no data or other information from the sample before the beginning of the OOS exercise will be allowed to pollute the recorded outcomes of the recursive prediction exercise.

These results largely confirm the findings reported in Table 3. Indeed, the GARCH that always includes the (squared) tone variable and the one that adds it only when it is negative, they both outperform the standard GARCH(1,1) model. Interestingly, the two augmented models seem to display a similar forecasting performance; however, this is not surprising considering that the tone is on average negative and therefore the two models tend to be often equivalent in terms of information they are based on. Indeed, as showed in Table 3, the sum of γ and θ in the GARCHAND model is equal to 1.23, which is close to the estimate of the γ in the GARCHX model (which equals 1.19). Therefore, we may conclude that a simple GARCH with an exogenous variable capturing the tone of the media coverage is useful to predict the conditional variance of FTSE 100 (CtO) returns.

4.3 Robustness checks

To test the robustness of our results, we also use our proposed news-augmented models to forecast the conditional variance of the CtO returns on the FTSE100. Notably, these

[12] Obviously, as we are implementing the exercise from the point of view of a trader / risk manager willing to exploit the forecast in real time, we only focus on the framework that uses the lag of the news-related variable to forecasts the variance at time $t + 1$.

returns are much less volatile than the OtC ones. In this framework, the news-related variable measured during the day t is used to forecast the variance arising overnight, i.e., between the closing of day t and the opening of day $t+1$. Table 6 shows the estimates for the already selected four best augmented models and for the GARCH(1,1). All the four models, including the one based on media attention (which is however included only when the volatility in the previous period was higher than 2%) show a better in-sample fit than the standard GARCH(1,1) model. In addition, the estimated γ coefficients are always statistically significant at 5% or lower, with the exception of the GARCHAND model based on the tone variable, where the estimates of γ and θ are characterized by a large standard error, probably due to an identification problem. Indeed, as discussed below, the tone variable is steadily negative and therefore the GARCHX and GARCHAND are overlapping.

Table 7 reports the OOS predictive performance of the four models and the standard GARCH(1,1) for the conditional variance of the FTSE100 CtO returns. We notice that all the selected models outperform the GARCH(1,1); indeed, at least when a less penalizing absolute value function is used, the DM test always leads to a rejection of the null hypothesis of equal forecasting accuracy. In terms of both RMSE and MAE, the GARCHX and GARCHND model based on the tone variables outperform all the others. However, in contrast with the results obtained for the OtC returns, in the case of overnight variance the growth in the number of articles seems to show some predictive power, even if less than the tone variable.

5. Conclusion

In this paper we have proposed a set of alternative augmented GARCH models that include variables that measure the coverage and the tone of the press coverage of a specific event, the Brexit referendum of 2016 and its subsequent developments. Such an event deemed to have had a great impact on the decision of stock market operators, thus impacting stock market volatility. We have assessed the predictive performance of these models comparing them with a standard GARCH model. With respect to the existing literature, we have studied the increases in predictive accuracy due to the inclusion of sentiment/attention indices specifically related to a single socio-political issue, such as the exit process of the UK from the European Union. Differently from us, previous research tended to focus alternatively on: 1) the general tone and level of media coverage reserved for the markets as a whole, unrelated to any specific political or macroeconomic event in progress; 2) the tone and level of firm-specific media coverage, referring to individual listed companies, or elements such as the deviation of results from consensus. In this sense, our proposed approach, even if limited to studying the volatility of stock returns, represents a compromise between a "generalist" and a "specific" approach. As far as the construction of the media attention and tone variables is concerned, to the best of our knowledge we are the first to the extensive information coming from GDELT to forecast stock return volatility.

We find that the innovative models that include the media coverage tone series outperform GARCH in three different specifications: GARCHX, GARCHND with κ = 10% and GARCHAND. The only model based on the series of growth rate of the number of articles that ensures a higher performance than GARCH is GARCHND with κ = 20%. From these results it emerges that knowing the tone with which the author reports the news significantly improves the forecast of the variance when it is always considered (GARCHX), when it is only considered for a certain level of previously predicted variance (GARCHND) and if an asymmetric effect dictated by the negative sign of the exogenous variable (GARCHAND) is implemented. The growth rate in the number of articles is only significant in high variance contexts (GARCHND).

A number of further extensions of this research would be possible. As far as the GARCHND model is concerned, the dummy variable is only activated for conditions that refer to the outputs of the model. However, it would be interesting to activate the dummies based on the levels of effective realized variance in the previous day, immediately measurable after the closing of the markets. This would mean linking the activation of the exogenous variable to the same measure chosen as a proxy for the variable to be predicted. Moreover, it would be interesting to implement both a "threshold of variance" activation effect and an asymmetric effect at the same time. We decided to test the two hypotheses separately for the sake of simplicity, but we do not exclude that they may occur together and possibly improve the prediction compared to a GARCH model (1,1).

As far as the GARCHAND models are concerned, it may be interesting to evaluate their performance compared to more consolidated asymmetric models such as the GJR-

GARCH, following Blasco, Corredor and Santamaria (2005). If the GARCHAND models result in being superior to some more successful asymmetric models, we could speak more properly of a real asymmetric effect due to the sign of the exogenous variable and not to that of returns. Another interesting test, following Antweiler and Frank (2004), may concern the study of a *news impact curve* estimated for the impact of the chosen exogenous variables on the variance of returns.

Finally, it might be interesting to study the simultaneous inclusion of both series based on press news, in order to alternatively test whether they contain information independent of each other (as their zero linear correlation appears to suggest), worsen the performance of the GARCH model, or even, surprisingly, improve the results of the models developed. Moreover, starting from the results in Table 7 regarding the significance of the first lag of the series of media coverage tone in the GARCHX and GARCHAND models, it may prove informative to test the effects of the inclusion of both the current and the one-day lagged value. In this sense, this would represent a genuine test of an actual "echo" effect for the tone with which the news are reported.

References

Andrei, D. and Hasler, M. (2014). Investor attention and stock market volatility. *Review of Financial Studies*, 28, 33–72.

Antweiler, W., and Frank, M. Z. (2004). Is all that talk just noise? The information content of Internet stock message boards. *Journal of Finance*, 59, 1259–94.

Belke, A., Dubova, I., and Osowski, T. (2018). Policy uncertainty and international financial markets: The case of Brexit. *Applied Economics*, 50, 1-19.

Berry, T.D., and Howe, K.M. (1994). Public information arrival. *Journal of Finance*, 49, 1331-1346.

Blasco, N., Corredor, P., Del Rio, C., and Santamaría, R. (2005). Bad news and Dow Jones make the Spanish stocks go round. *European Journal of Operational Research*, 163, 253-275.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31, 307-327.

Bollerslev, T., and Wooldridge, J. M. (1992). Quasi-maximum likelihood estimation and inference in dynamic models with time-varying covariances. *Econometric Reviews*, 11, 143-172.

Brown, G. W., and Cliff, M. T. (2005). Investor sentiment and asset valuation. *Journal of Business*, 78, 405-440.

Campbell, J. Y., and Hentschel, L. (1992). No news is good news: An asymmetric model of changing volatility in stock returns. *Journal of Financial Economics*, 31, 281-318.

Chen, H., De, P., Hu, Y. J., and Hwang, B. H. (2014). Wisdom of crowds: The value of stock opinions transmitted through social media. *Review of Financial Studies*, 27, 1367-1403.

Da, Z., Engelberg, J., and Gao, P., (2011). In search of attention. *Journal of Finance*, 66, 1461-1499.

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

Diebold, F. X., and Mariano, R. S. (1995). Comparing predictive accuracy. *Journal of Business and Economic Statistics*, 20, 134-144.

Dimpfl, T. and Jank, S. (2016). Can internet search queries help to predict stock market volatility? *European Financial Management*, 22, 171–192.

Engle, R. (2002). New frontiers for ARCH models. *Journal of Applied Econometrics*, 17, 425-446.

Fang, L.H., and Peress, J. (2009). Media coverage and the cross-section of stock returns. *Journal of Finance*, 64, 2023-2052.

Glosten, L. R., Jagannathan, R., and Runkle, D. E. (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. *Journal of Finance*, 48, 1779-1801.

Goddard, J., Kita, A. and Wang, Q. (2015). Investor attention and FX market volatility. *Journal of International Financial Markets, Institutions Money*, 38, 79–96.

Guidolin, M., Orlov, A. G., and Pedio, M. (2017). How good can heuristic-based forecasts be? A comparative performance of econometric and heuristic models for UK and US asset returns. *Quantitative Finance*, 18, 139-169.

Hansen, P. R., and Lunde, A. (2005). A forecast comparison of volatility models: Does anything beat a GARCH (1, 1)? *Journal of Applied Econometrics*, 20, 873-889.

Kierzenkowski, R., Pain, N., Rusticelli, E., and Zwart, S. (2016). The economic consequences of Brexit: A taxing decision. *OECD Economic Policy Papers*, 16, 1.

Mitchell, M.L., and Mulherin, J.H. (1994). The impact of public information on the stock market. *Journal of Finance*, 49, 923--950.

Raddant, M. and Karimi, F. (2016). Cascades in real interbank markets. *Computational Economics,* 47, 49-66.

Shen, D., Urquhart, A., and Wang, P. (2019). Does twitter predict Bitcoin? *Economics Letters*, 174, 118-122.

Tetlock, P. C. (2007). Giving content to investor sentiment: The role of media in the stock market. *Journal of Finance*, 62, 1139-1168.

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

Wohlfarth, P. (2018). Measuring the impact of monetary policy attention on global asset volatility using search data. *Economics Letters*, 173, 15-18.

Table 1

Descriptive Statistics

The table shows the main descriptive statistics of the *Open-to-Close* (OtC) and *Close-to-Open* (CtO) FTSE100 daily returns and of their realized volatility, as well as of the news-related variables. The realized volatility is expressed in annualized terms assuming 252 trading days per year.

	Returns				Exogenous Variables		
	Open-to-Close	Close-to-Open	Realized Volatility (OtC)	Realized Volatility (CtO)	Tone	Art Growth	Articles
Mean	0.03%	-0.01%	11.34%	1.30%	-1.46	11.35%	683
Stadard Deviation	0.82%	0.12%	7.59%	1.39%	1.05	61.45%	981
Kurtosis	1.99	9.30	146.11	21.98	13.73	25.69	149.09
Skewness	-0.04	-0.71	8.73	3.40	-1.82	3.67	9.88
Minimum	-3.68%	-0.99%	1.83%	0.00%	-11.76	-100.00%	0
Maximum	3.14%	0.63%	150.53	15.67%	1.81	475.00%	18215
Sample Size	718	781	781	781	781	781	781

Table 2

Sample Correlations

The table shows the pair-wise correlations between the realized variance (volatility) of the FTSE100 OtC daily returns and the news related series used in the analysis (as well as their squares). Boldfaced sample estimates are significant at a test size of 5% or lower.

	Realized Variance	Realized Volatility	Tone	Tone^2	Art Growth	Art Growth^2	Articles
Realized Variance	**1.00**	-	-	-	-	-	-
Realized Volatility	**0.81**	**1.00**	-	-	-	-	-
Tone	-0.03	-0.11	**1.00**	-	-	-	-
Tone^2	0.03	0.11	**-0.79**	**1.00**	-	-	-
Art Growth	**0.23**	0.20	-0.03	0.06	**1.00**	-	-
Art Growth^2	**0.27**	**0.26**	0.03	0.00	**0.73**	**1.00**	-
Articles	**0.63**	**0.46**	-0.01	-0.03	**0.23**	0.18	**1.00**

Table 3

In-Sample Results for News-Augmented Conditional Heteroskedastic Models

The table shows the results of the in-sample estimation of the GARCH model and its news-augmented versions. The significance of the individual coefficients is indicated according to the p-values of a standard (Wald) *t*-test for size levels equal to 1%, 5%, and 10% (***, ** and *, respectively). In-sample estimation statistics that are superior to a standard GARCH model are boldfaced. The coefficient estimates that are significant at a level of 5% or lower are also boldfaced.

	In-sample Estimation Statistics			Coefficients				
	RMSE	MAE	MAPE	ω	α	β	γ	θ
GARCH	0.03431	0.00526	1.04813	0.00029	0.11079	0.84742		
				(0.05283)	(0.01870)	(0.01408)		
GARCHX Tone	**0.03430**	**0.00525**	1.05403	0.00023	**0.11825*** **	**0.78381*** **	**1.18743*** **	
				(0.05483)	(0.02757)	(0.03867)	(0.46141)	
GARCHX Art Growth	0.03431	0.00530	1.04784	0.00029	**0.10764*** **	**0.84794*** **	0.00000	
				(0.05581)	(0.01868)	(0.01550)	(0.01002)	
GARCHND Tone (κ=10%)	0.00346	0.00568	1.28394	0.00276	**0.28496*** **	**0.18133*** **	**3.22938*** **	
				(0.03951)	(0.05637)	(0.03024)	(1.06214)	
GARCHND Tone (κ=20%)	0.03489	0.00639	1.20261	0.00012	**0.21963*** **	**0.56197*** **	**0.12221** **	
				(0.04161)	(0.05038)	(0.03107)	(0.04714)	
GARCHND Art Growth (κ=10%)	0.00343	0.00527	1.03748	0.00030	**0.10331*** **	**0.84299*** **	0.01441	
				(0.05540)	(0.19349)	(0.01388)	(0.01547)	
GARCHND Art Growth (κ=20%)	0.03432	0.00533	1.05191	0.00031	**0.11065*** **	**0.84242*** **	3.28805	
				(0.05220)	(0.01970)	(0.01455)	(27.5701)	
GARCHAND Tone	**0.03430**	**0.00524**	1.05194	0.00023	**0.11736*** **	**0.78655*** **	0.79700	0.44065
				(0.05518)	(0.02721)	(0.03879)	(3.54164)	(6.25922)
GARCHAND Art Growth	0.03431	0.00530	1.04708	0.00029	**0.10837*** **	**0.84631*** **	0.00000	
				(0.05367)	(0.01879)	(0.01520)	(0.01103)	

Table 4

In-Sample Results for Augmented Conditional Heteroskedastic Models Based on Contemporaneous News-related Variables

The table shows the results of the in-sample estimation of the GARCH model and its augmented versions when the news-related variables at time $t + 1$ are used to forecast the conditional variance at time t. The significance of individual coefficients is indicated according to the p-values of a standard (Wald) t-test for size levels equal to 1%, 5%, and 10% (***, ** and * respectively). In-sample fitting results which are superior to a standard GARCH model are boldfaced. The coefficient estimates that are significant at a level of 5% or lower are also boldfaced.

	RMSE	MAE	MAPE	ω	α	β	γ	θ
GARCH	0.03431	0.00530	1.04710	0.00029 (0.05548)	**0.10726***** **(0.01857)**	**0.84850***** **(0.01398)**		
GARCHX Tone	**0.03430**	**0.00523**	**1.03890**	0.00020 (0.05402)	**0.11165***** **(0.02503)**	**0.81322***** **(0.03334)**	**0.82390**** **(0.37699)**	
GARCHX Art Growth	**0.03420**	**0.00527**	**1.03985**	0.00032 (0.05594)	**0.11521***** **(0.02143)**	**0.82256***** **(0.01873)**	**2.16316*** **(1.26814)**	
GARCHND Tone ($\kappa = 10\%$)	**0.03430**	**0.00517**	**0.99793**	0.00030 (0.06453)	**0.09094***** **(0.01850)**	**0.82573***** **(0.01511)**	**1.01568***** **(0.28396)**	
GARCHND Tone ($\kappa = 20\%$)	0.03431	0.00534	1.05569	0.00034 (0.05435)	**0.11442***** **(0.02108)**	**0.83134***** **(0.01547)**	0.53839 (1.11477)	
GARCHND Art Growth ($\kappa = 10\%$)	**0.03417**	**0.00519**	**0.98047**	0.00028 (0.05726)	**0.09473***** **(0.01289)**	**0.84428***** **(0.00841)**	**3.29909**** **(1.80058)**	
GARCHND Art Growth ($\kappa = 20\%$)	**0.03411**	0.00530	**1.04716**	0.00033 (0.05088)	**0.10758***** **(0.01934)**	**0.83989***** **(0.01475)**	4.30014 (9.36667)	
GARCHAND Tone	**0.03430**	**0.00523**	**1.03730**	0.00021 (0.05370)	**0.11100***** **(0.02481)**	**0.81526***** **(0.03359)**	0.68847 (4.32117)	0.15219 (7.11584)
GARCHAND Art Growth	**0.03420**	**0.00527**	**1.03746**	0.00032 (0.05580)	**0.11307***** **(0.02100)**	**0.82581***** **(0.01791)**	**2.28466**** **(1.33139)**	

Table 5

Out-of-Sample Predictive Performance for News-Augmented Conditional Heteroskedastic Models

The Table reports the OOS forecasting accuracy measures for alternative augmented GARCH models of the conditional variance of the (OtC) returns on the FTSE100 Index. The RMAPE measures the relative performance of the news-augmented model vs. the GARCH(1,1) model. A RMAPE lower than 100% implies that the augmented model outperforms the GARCH(1,1). The Theil's U compares the augmented model vs. a naïve benchmark that imposes the forecasted variance at time t+1 to equal the realized variance at time t; a value lower than 100% implies that the augmented GARCH model outperforms the benchmark. The Diebold-Mariano statistics for the null of equal predictive accuracy of the augmented GARCH model and the standard GARCH(1,1) are computed under two loss functions, i.e., the quadratic loss (DM SQ) and absolute value (DM AV). Significance at the size levels of 1%, 5%, and 10% is denoted by ***, **, and *, respectively. All the statistics that denote a forecasting performance superior to the GARCH benchmark are boldfaced.

	Out-of-sample Performance Statistics						
	RMSE	MAE	MAPE	RMAPE	Theil's U	DM SQ	DM AV
GARCHX Tone	**0.00468**	**0.00280**	**1.1634**	**76.79%**	90.49%	-0.70629	**-3.34435*** **
GARCHND Tone (κ=10%)	0.00506	0.00328	**1.4250**	**94.06%**	94.21%	1.85928	0.31903
GARCHND Art Growth (κ=20%)	0.00477	0.00322	**1.5110**	**99.73%**	96.48%	0.70734	0.31903
GARCHAND Tone	**0.00467**	**0.00280**	**1.1693**	**77.18%**	90.53%	0.72656	**-3.38672*** **

Table 6

In-Sample Results for News-Augmented Conditional Heteroskedastic Models for FTSE100 Close-to-Open Returns

The table shows the results of the in-sample estimation of the GARCH model and its news-augmented versions when the news-related variables at time t are used to forecast the conditional variance overnight (i.e., between the closing of time . The significance of individual coefficients is indicated according to the p-values of a standard (Wald) t-test for size levels equal to 1%, 5%, and 10% (***, ** and * respectively). In-sample fitting results which are superior to a standard GARCH model are boldfaced. The coefficient estimates that are significant at a level of 5% or lower are also boldfaced.

	In-sample Estimation Statistics		Coefficients				
	RMSE	MAE	ω	α	β	γ	θ
GARCH	0.03101	0.01459	0.00014	**0.03584*****	**0.95435*****		
			(0.00021)	(0.00517)	(0.00424)		
GARCHX Tone	**0.03099**	**0.01448**	0.00002	**0.02588*****	**0.96343*****	**0.34118*****	
			(0.00024)	(0.00616)	(0.00493)	(0.13150)	
GARCHND Tone (κ=1%)	**0.03099**	**0.01448**	0.00002	**0.28496*****	**0.18133*****	**3.22938*****	
			(0.00024)	(0.00616)	(0.00493)	(0.13153)	
GARCHND Art Growth (κ=2%)	**0.031**	**0.0145**	0.00008	**0.03319*****	**0.95724*****	**1.32250****	
			(0.00021)	(0.00478)	(0.00421)	(0.71203)	
GARCHAND Tone	**0.03099**	**0.01449**	0.00002	**002640*****	**0.96326*****	0.22236	0.55779
			(0.00024)	(0.00621	(0.00496)	(4.35084)	(30.6075)

Table 7

Recursive Predictive Performance of Alternative Models for the Variance of the FTSE100 Close-to-Open Index Returns

The Table reports OOS forecasting accuracy measures for alternative augmented GARCH models for the conditional variance of the (CtO) returns of the FTSE100 Index. The Theil's U compares the augmented model vs. a naïve benchmark that impose the forecasted variance at time t+1 equal to the realized variance at time t; a value lower than 100% implies that the augmented GARCH model outperforms the benchmark. The Diebold-Mariano statistics for the null of equal predictive accuracy of the augmented GARCH model and the standard GARCH(1,1) are computed under two loss functions, i.e., the quadratic loss (DM SQ) and absolute value (DM AV). Significance at the size levels of 1%, 5%, and 10% is denoted by ***, **, and *, respectively. All the statistics that denote a forecasting performance superior to the GARCH benchmark are boldfaced.

	Out-of-sample Performance Statistics				
	RMSE	MAE	Theil's U	DM SQ	DM AV
GARCH	0.11526	0.03667	49.28%		
GARCHX Tone	**0.11503**	**0.03431**	50.70%	-1.07646	**-4.79189*****
GARCHND Tone (κ=1%)	**0.11503**	**0.0343**	50.70%	-1.07611	**-4.79317*****
GARCHND Art Growth (κ=2%)	**0.11515**	**0.0364**	**47.89%**	**-1.62175***	**-2.72156*****
GARCHAND Tone	**0.11504**	**0.03444**	51.37%	-1.04650	**-4.79366*****

Figure 1

News related series

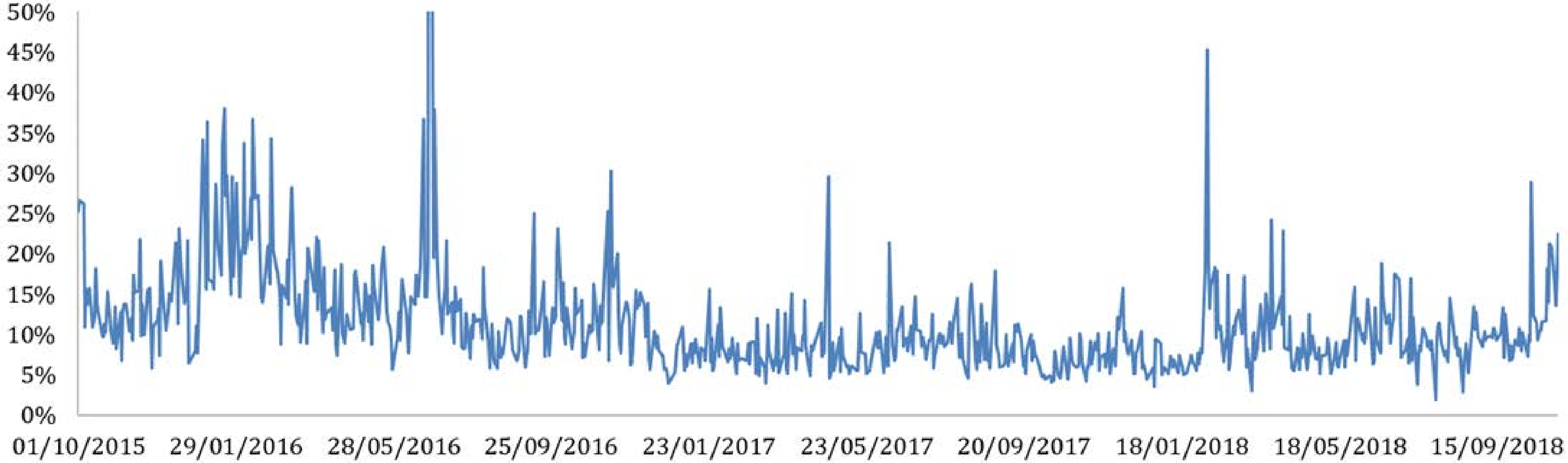


Figure 2

Growth Rate of the Number of Articles About "Brexit"

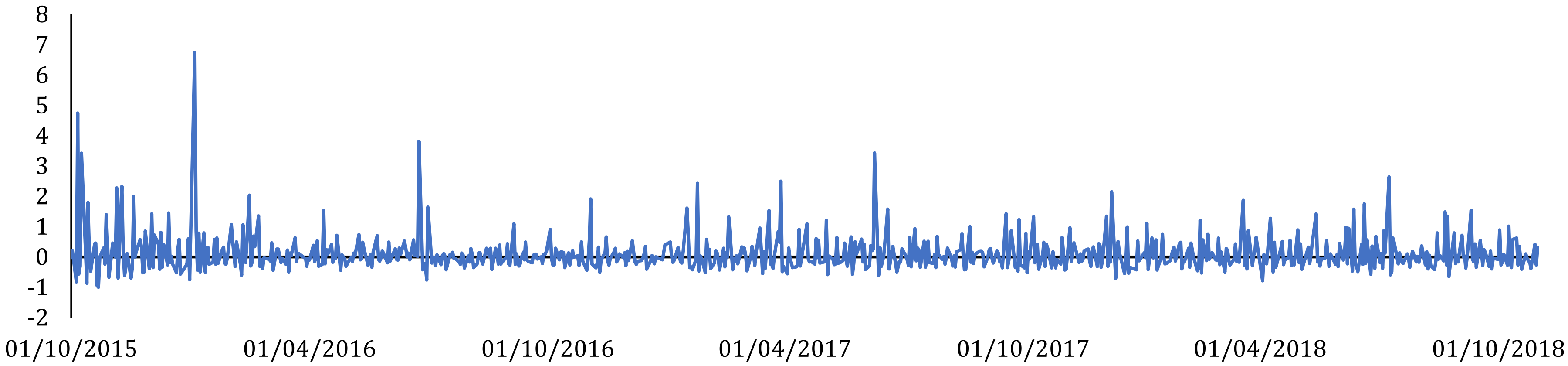


Figure 3

Tone of Media Coverage Regarding "Brexit"

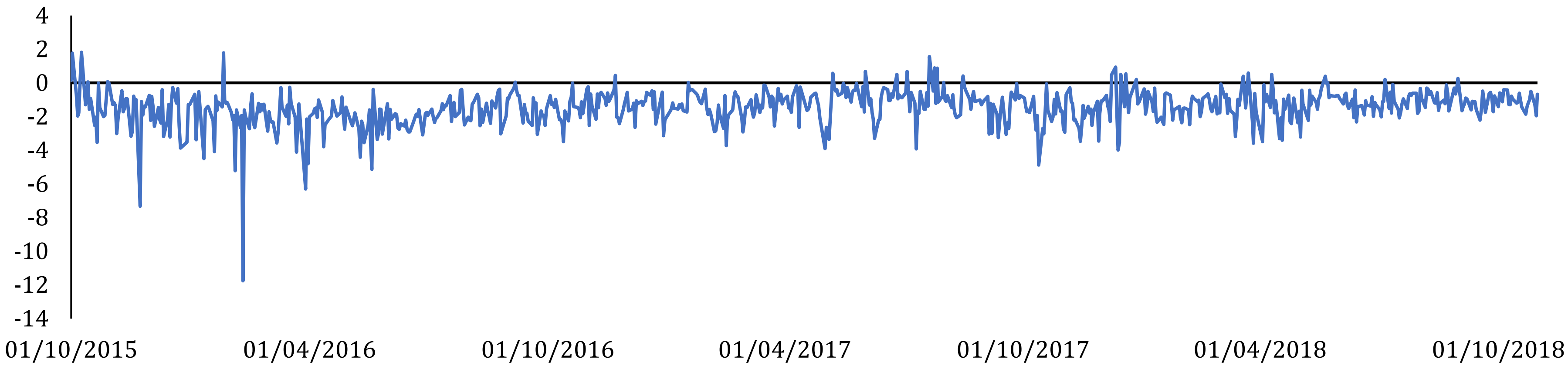


Figure 4

Realized Volatility of the *Open-to-Close* Returns on the FTSE100 Index

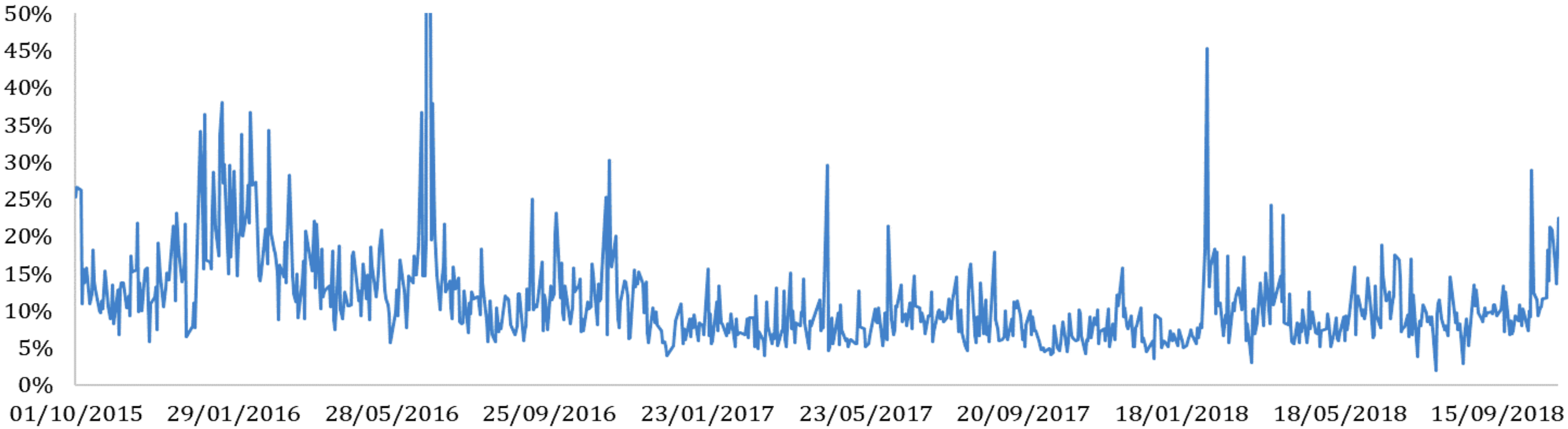
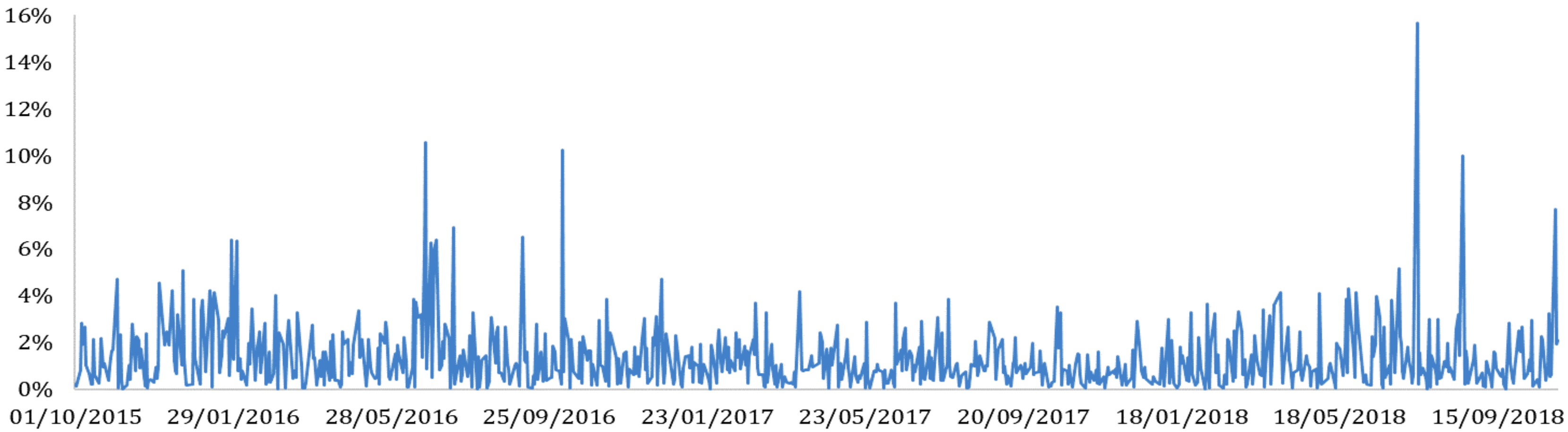


Figure 5

Realized Volatility of the *Close-to-Open* Returns of the FTSE100 Index



Appendix

The Global Knowledge Graph data file is the main source for our analysis. This output is organized around "*namesets*". A *nameset* is a set of names (of people, such as politicians or organizations and companies), locations (such as UK, London, Italy, etc.) and themes that tend to appear together in the articles. An example of "*nameset*" may be $\{David\ Cameron, Boris\ Johnson, United\ Kindom, European\ Union, ELECTION\}$ where ELECTION is one of the themes covered by the GDELT Project. All the articles that contain a mention (or more) to all the words contained by the *nameset* will be grouped together into one record.

A query through the GKG tool using a certain keyword, such as "Brexit", will return all the "*namesets*" that contain the keyword and the information associated to each "*nameset*". Therefore, each record in the dataset (i.e., each row in the output) represents a "*nameset*". While many fields are potentially available, in Table A we report a sample of the extraction that we actually used in our analysis. Below we provide a short description of the fields involved.

- DATE: it is the date (in YYYYMMDD format) on which the news media used to construct the file was firstly published.
- NUMARTS: it is the number of articles that contain one or more mentions to the "*nameset*" and that have been therefore grouped together
- THEMES: it is the list of themes mentioned in the *nameset*; the themes are from a list prepared by the GDELT team to categorize the articles.
- LOCATIONS: it is the list of locations mentioned in the *nameset*.
- PERSONS: it is the list of persons mentioned in the *nameset*.
- ORGANIZATIONS: it is the list of the organizations or companies mentioned in the *nameset*.
- TONE: it is the average tone of the articles containing one (or more) mention of the *nameset*. The score ranges between -100 (extremely negative) and +100 (extremely positive); it is computed as positive minus negative score.
- POSITIVE SCORE: it is the percentage (from 0 to 100) of the words in the article that are found to have a positive emotional connotation.
- NEGATIVE SCORE: it is the percentage (from 0 to 100) of the words in the article that are found to have a negative emotional connotation.
- POLARITY: it is the percentage of the words in the article that are fund to have an emotional connotation (either negative or positive).
- SOURCES: it contains the list of all the articles with at least one mention to the *nameset*.
- SOURCEURLS: it contains the URLs of all the articles with at least a mention to the *nameset*.

Note that the GDELT project is in continuous evolution. These indications refer to the time at which we performed the extraction, i.e., November 2018.

Table A – Sample Dataset

Date	NumAr	Themes	Locations	Persons	Organizations	ToneData	PositiveSc	NegativeSc	Polarity	Sources	SourceUI s
01/10/2015	8	TAX_FNCACT;TAX_FNCACT_M	1#China#CH#	margaret thatcher;nigel far	united kingdom in	0.42826552	3.64025696	3.21199143	6.85224839	adelaidenow.	http://www.adelaider
01/10/2015	1	CRISISLEX_CRISISLEXREC;TA	4#Doncaster,	nigel farage;george osborn	european union;y	-0.48543689	3.15533981	3.6407767	6.7961165	telegraph.co.	http://www.telegraph
01/10/2015	1	TAX_FNCACT;TAX_FNCACT_M	4#London, Lc	michael holden;margaret tl	party for britain;u	-1.72413793	2.01149425	3.73563218	5.74712644	breitbart.com	http://www.breitbart
02/10/2015	1	TAX_FNCACT;TAX_FNCACT_M	1#United Kin	margaret thatcher;david ca	european union;re	0	2.91262136	2.91262136	5.82524272	latestnewslin	http://latestnewslink.
02/10/2015	1	TAX_FNCACT;TAX_FNCACT_M	1#United Kin	margaret thatcher;david ca	european union	1.34529148	2.69058296	1.34529148	4.03587444	in-cyprus.con	http://in-cyprus.com/
02/10/2015	1	TAX_FNCACT;TAX_FNCACT_M	1#United Kin	margaret thatcher;david ca	european union;e	1.92926045	3.8585209	1.92926045	5.78778135	breitbart.com	http://www.breitbart
02/10/2015	1	TAX_FNCACT;TAX_FNCACT_M	4#London, Lc	paul sandle;margaret thatc	european union;e	1.92307692	3.84615385	1.92307692	5.76923077	dailymail.co.u	http://www.dailymail
02/10/2015	1	TAX_FNCACT;TAX_FNCACT_M	4#London, Lc	boris johnson		0	3.76344086	3.76344086	7.52688172	wiltshirebusi	http://www.wiltshire
02/10/2015	1	USPEC_POLICY1;GENERAL_G(	4#Belfast, Bel	david phinnemore;cathal m	ulster university u	0.498132	1.99252802	1.49439601	3.48692403	communityni	http://www.commun
02/10/2015	50	TAX_FNCACT;TAX_FNCACT_M	4#London, Lc	boris johnson		0.68965517	4.13793103	3.44827586	7.5862069	dailyecho.co.	http://www.dailyech
02/10/2015	1	TAX_FNCACT;TAX_FNCACT_M	1#United Kin	margaret thatcher;david ca	european union	0.87336245	3.05676856	2.18340611	5.24017467	todayszaman	http://www.todaysza
02/10/2015	1	TAX_FNCACT;TAX_FNCACT_M	4#London, Lc	nigel farage;david cameron	united kingdom in	-0.18587361	2.41635688	2.60223048	5.01858736	mundinews.c	http://mundinews.co
02/10/2015	1	TAX_FNCACT;TAX_FNCACT_C	1#France#FR	nigel farage;alan walters;m	european exchang	-1.14942529	2.80970626	3.95913155	6.7688378	dailymail.co.u	http://www.dailymail