News Augmented GARCH(1,1) Model for Volatility Prediction

ZRYAN A. SADIK[†], PARESH M. DATE[†] AND GAUTAM MITRA[††]

[†]*Department of Mathematics, Brunel University London,*
Uxbridge, London, UB8 3PH, UK.
(zryan.sadik, paresh.date)@brunel.ac.uk

[††]*OptiRisk Systems,*
OptiRisk R&D House,
One Oxford Road,
Uxbridge, London, UB9 4DA, UK.
gautam@optirisk-systems.com

[†]*Corresponding author: zryan.sadik@brunel.ac.uk*

Forecasting the volatility of stock return plays an important role in the financial markets. The GARCH model is one of the most common models used for predicting asset price volatility from the return time series. In this study, we have considered quantified news sentiment and its impact on the movement of asset prices as a second source of information, which is used together with the asset time series data to predict the volatility of asset price returns. We call this NA-GARCH (news augmented GARCH) model. Our empirical investigation compares volatility prediction of returns of 12 different stocks (from two different stock markets), with 9 data sets for each stock. Our results demonstrate that NA-GARCH provides a superior prediction of volatility than the "plain vanilla" GARCH and EGARCH models. These results vindicate some recent findings regarding the utility of news sentiment as a predictor of volatility, and also vindicate the utility of our novel model structure combining the proxies for past news sentiments and the past asset price returns.

Keywords: volatility prediction, GARCH, EGARCH, news sentiment, news impact scores.

1. Introduction and Background

Volatility forecasting of stock return plays an important role in numerous financial applications in the financial markets. From the beginning of the 21st century, a number of researchers have analysed how news and market sentiment influence the financial markets and its participants. The researchers arrived at different conclusions. For instance, Mitchell & Mulherin (1994) observed that the relation between news and market activity is not particularly strong and the patterns in news announcements do not explain the day-of-the-week seasonalities in market activity. On the other hand, Kalev *et al.* (2004) gave evidence that news arrivals display a very strong pattern of autocorrelation. In Engle & Ng (1993), the authors defined the news impact curve that measures how new information is incorporated into volatility estimates. Their results suggest that the EGARCH model, first proposed in Nelson (1991), can capture most of the asymmetry without the need of modelling news impact independently. However, they report evidence that the variability of conditional variance implied by the EGARCH model is too high. The early research of applying news analysis to financial markets focused on equities. Later, macroeconomic news and its impact on fixed income has been studied extensively. In the domain of fixed income macroeconomic announcements (news data) also influence asset price. For example, Arshanapalli *et al.* (2006) investigated the effects of macroeconomic news on time-varying volatility as

well as time-varying covariance for the US stock and bond markets; they found that stocks and bonds have higher volatility on the day of macroeconomic announcements. Li & Engle (1998) studied the effect of the macroeconomic public announcements and revealed the heterogeneous persistence from scheduled news versus non-scheduled news. They examined the reaction of the Treasury futures market to the periodically scheduled announcements of prominent U.S. macroeconomic data.

More recent focus of relevant research is the dynamic relationship between news sentiment and the changes in the asset price dynamics. Tetlock (2010) measured public information using firms' stock returns on news days in the Dow Jones archive. He found four patterns in post-news returns and trading volume that are consistent with the asymmetric information model's predictions. Ho *et al.* (2013) compared macroeconomic news sentiment with firm-specific news sentiment, they found that the latter accounts for a greater proportion of overall volatility persistence. Crouhy & Rockinger (1997) confirmed that volatility rises more in response to bad news than to good news. Riordan *et al.* (2013) confirmed that negative news messages induce stronger market reactions than positive ones. The empirical results of Song (2010) indicated that unexpected bad news about a particular portfolio tends to increase the volatility of the returns on other correlated portfolios, whereas unexpected good news about a particular portfolio has an opposite impact on the volatility of correlated portfolios. Chen & Ghysels (2011) found that moderately good (intra-daily) news reduces volatility (the next day), while both very good news (unusual high intra-daily positive returns) and bad news (negative returns) increase volatility, with the latter having a more severe impact. In addition, Sidorov *et al.* (2013) considered trading volume as a proportional proxy for information arrivals to the market and the daily number of press releases on a stock (news intensity) as an alternative explanatory variable in the basic equation of GARCH model. They showed that the GARCH(1,1) model augmented with volume does remove GARCH and ARCH effects for the most of the companies, while the GARCH(1,1) model augmented with news intensity has difficulties in removing the impact of log return on volatility. Later in 2014, Sidorov *et al.* (2014) analyzed the impact of news intensity as extraneous sources of information on stock volatility. Their results showed that the GARCH(1,1) model augmented with the news intensity performs better than the pure GARCH model.

There is a strong, yet complex relationship between market sentiment and news. Traders and other market participants digest news rapidly and update their asset positions accordingly. However, for models to incorporate news directly and automatically, we require quantitative inputs, whereas raw news is qualitative data. Companies such as RavenPack and Thomson Reuters have developed linguistic analytics which process the textual input of news stories to determine quantitative sentiment scores. Both sets of RavenPack and Thomson Reuters data are similar in structure (see RavenPack (2014) and Thomson Reuters (2010)). However, Ravenpack data is used in this study since this was the only data source available to authors from their commercial sponsor (see Mitra & Mitra (2011) for a detailed anaylsis of RavenPack data). As compared to the amount of effort expended in forecasting volatility from return time series alone, the academic literature on exploiting these sentiment scores for pricing or forecasting seems to be somewhat limited. For example, Tetlock (2007) explored the interactions between investor sentiment and stock market, Mitra *et al.* (2009) used quantified news and implied volatility to improve risk estimates as the market sentiment and environment changes.

In our work presented here, we use RavenPack's news sentiment score as a quantitative proxy for news sentiment. This proxy is used as an exogenous term in a modified version of GARCH model proposed by Bollerslev (1986). We propose a new model structure to introduce the impact of news into volatility prediction in a way which is meaningful from economic point of view. While other GARCH models exist (see, e.g. EGARCH and GED-GARCH by Nelson (1991), GRS-GARCH by Gray (1996), GJR-GARCH by Glosten *et al.* (1993) and semi-parametric extension SEMIFAR-GARCH by Feng *et*

al. (2007), among others), we have tested variants of GARCH models, via GARCH-t (GARCH with t-distributed residuals) and TGARCH (threshold GARCH proposed by Zakoian (1994)) in conjunction with our news proxy. These models were calibrated and tested in terms of volatility forecasting and risk prediction ability, on datasets from two different financial markets. The structure of our proposed model in this study is based on GARCH(1,1) model. However, the proposed model could be easily extended and modified to account for more general GARCH(p,q) models, with $\max(p,q) > 1$. The simple GARCH(1,1) model has been found to adequately fit many economic and financial time series as well as proven surprisingly successful in predicting conditional variances; see for example, Sharma et al.(1996); Hsieh(1988,1989); Bollerslev(1987); Baillie and Bollerslev(1989) and McCurdy and Morgan(1987). Further, there is evidence in the literature that it is hard to beat GARCH(1,1) in terms of its forecasting ability (Hansen & Lunde (2005)). This was also confirmed in our experiments on news enhanced versions of GARCH, GARCH-t and T-GARCH models. Therefore, the use of higher order GARCH model has not been reported in our study. While presenting the results of our study, we have restricted our focus on the simple GARCH(1,1) model and its news augmented version. As mentioned earlier, literature provides some evidence of EGARCH model to capture asymmetry in volatility which may result from differing impact of positive and negative news (Engle & Ng (1993)). Hence we have compared both these models with each other as well as the EGARCH(1,1) model on multiple datasets.

The structure of our NA-GARCH model is novel and our study vindicates the findings of other researchers, namely, Mitra *et al.* (2009) and Arbex-Valle *et al.* (2013) who have used factor models as predictors of realized volatility. The broad conclusion of the earlier studies which is reinforced by our study is that in the financial markets the use of news sentiment leads to better predictions of the volatility of asset returns.

The rest of the paper is organized as follows. In section 2 we present the granularity of the experimental data and we also explain the two streams of time series, namely, the market data and news metadata that are used for numerical experiments. In section 3 we describe the model using which the stream of sentiment meta data is turned into impact of news items on asset price: this we call the news impact model. We describe the classical GARCH and EGARCH models as well as the new NA-GARCH model. In section 4 we describe the issues of model calibration, model fitting and the performance measures. In section 5 we analyse the computational results of the empirical investigation of the two sets of six assets comparing the performances of the GARCH, EGARCH and NA-GARCH models using the chosen performance measures. In section 6 we set out our discussions and conclusions.

2. Data

Our experimental dataset comprises two streams of time series data: the daily market (price) data of the closing prices and the news metadata (supplied by RavenPack) for each asset considered in the experiment. The news meta data provides us quantified values of sentiment which in turn is processed into news impact scores; this news impact model is described in section 3.1.

2.1 *Data Granularity*

The trading day starts at 08:00 hours and ends at 16:30 hours. Thus, in a trading day the total number of minutes is 510. The frequency of the news impact scores aligned to the trading hours of 08:00-16:30. Therefore, any news arriving overnight or during the weekend is bucketed into the next morning or days first minute, where the size of a bucket is 1 minute. Hence, the assumption is taken that the impact of such overnight and weekend news is incorporated into prices the following day.

For this study, the stock market daily closing price data has been used. Thus, the news impact score was taken at the last minute in the trading day, which was the aggregation of all the news data from 08:00 to 16:30. The news impact scores actually represented by two data streams and they normalized or scaled in the range between 0 and 1 for positive impact, whereas between -1 and 0 for negative impact. Note that although news data is intraday, it is not conformed to a regular time scale thus there is not a guaranteed number of data points within each day for a particular asset.

2.2 *Market Data*

The time series data used for modelling volatility in this study is the stock market daily closing price of twelve assets from FTSE100 and EUROSTOXX50 (six assets from each index). We extracted the data of 12 assets across different sectors based on their cap-weights. The reason for our concern with the cap-weights is that we found companies with large market capitalisation have wide coverage of news which guarantees a sufficient number of news data points. The data covers seven different sectors which are Pharmaceuticals & Biotechnology, Insurance, Oil & Gas, banking, Mobile Telecommunications, Food & Beverage and Chemicals. Table 1 lists these sectors and each asset that is belonging to each sector. The data was obtained from Interactive Data.

	FTSE100			**EUROSTOXX50**	
	Asset	**Sector**		**Asset**	**Sector**
1	AstraZeneca	Pharmaceuticals & Biotechnology		Allianz	Insurance
2	Aviva	Insurance		Anheuser-Busch	Food & Beverage
3	BP	Oil & Gas		Banco Santander	Banking
4	GlaxoSmithKline	Pharmaceuticals & Biotechnology		Bayer	Chemical
5	Lloyds Bank	Banking		Deutsche Bank	Banking
6	Vodafone	Mobile Telecommunications		Total	Oil & Gas

Table 1. List of twelve assets from FTSE100 and EUROSTOXX50

The data for each asset covers daily closing prices from 3 January 2005 to 31 December 2015 (ten years). We chose a rolling window of size 750, the number of consecutive observations per rolling window, and increments of 253 observations between successive rolling windows such that each rolling window starts from the very beginning of January and represents roughly a period of three years. The data for each individual asset is divided into nine datasets such that each dataset contains 750 observations. The data used in model fitting are different from those used in predicting evaluation. Typically, we divided each dataset into two sub-periods. Considering a dataset consists of T observations, $p_1, \ldots, p_T$, we divide the data as $\{p_1, \ldots, p_n\}$ and $\{p_{n+1}, \ldots, p_T\}$ where n is the initial forecast origin. According to Tsay (2008), a reasonable choice for splitting is $n = 2T/3$, where T is the data points and n is the initial forecast origin. Therefore, since each dataset in this paper has 750 observations, we used the first 500 observations (two years) as the in-sample data for fitting the models in order to estimate the parameters of the models, while the remaining of 250 observations (one year) are taken as the out-of-sample data and used to evaluate the forecasting performance of the models.

The purpose of choosing ten years of data is to study the performance of our proposed model in different economic periods (including a recession period in 2008) and different markets (UK and in Europe).

In this study, daily returns (r_t) were calculated as the continuously compounded returns which are

the first difference in logarithm of closing prices of the asset of successive days:

$$r_t = \log\left(\frac{P_t}{P_{t-1}}\right) \tag{2.1}$$

where P_t and P_{t-1} are the daily closing price of the index at current and previous days, respectively. For all the datasets in this study, we utilise the Ljung-Box test to examines whether there is significant evidence for non-zero correlations at lags 1-20. Small p-values (*i.e.*, less than 0.05) suggest that the time series is stationary. The test shows that stationarity is satisfied for every dataset.

2.3 *News Metadata*

News analytics data is presented in a metadata format where news is given characteristics such as relevance, novelty and sentiment scores according to an individual asset. The analytical process of producing such scores is fully automated from collecting, extracting, aggregating, to categorising and scoring. The result is an individual score assigned to each news article for each characteristic using scales from $0 - 100$. A news sentiment score measures the emotional tone within a news item and varies between positive and negative. Sentiment score is a value falling within a range consisting of a minimum and maximum depicting the overall tone of a news article. Depending on the measurement of scale, the exact polarity of sentiment in the news can be deduced, *i.e.* Thomson Reuters assign probabilities to the moods "Positive", "Neutral" and "Negative" to infer an overall sentiment that is the average of all 3 scores (Thomson Reuters (2010)), whereas RavenPack directly produce a sentiment score belonging to the range 0-100 that then allows a conclusion of positivity or negativity, where 0 is the most negative score and 100 is the most positive score.

We have used RavenPack news metadata in all our experiments. The RavenPack data is represented in two separate files, the first contains equity (company) related analytics and the second one contains global macro analytics. Each record in equity news analytics file contains 46 fields including a time stamp, reference identifiers, scores for relevance, novelty and sentiment, and unique identifiers for each news story analyzed. The explanation of the descriptions and knowledge of RavenPack data fields are represented in the RavenPack news analytics manual (RavenPack(2014)).

3. Models

We introduce four models in this section; these are news impact model, plain vanilla GARCH model, EGARCH model and news augmented GARCH model.

3.1 *News Impact Model*

As mentioned in the previous section (2.3), we have used RavenPack's news analytics database in our research. In order to derive suitable news impact terms (positive and negative impact scores) and utilize them as proxies of good and bad news, the following two points have to be taken in consideration:

1. An expression has to be found which describes the attenuation of the news sentiment score.

2. The impact of a news item will decay over the time and older news has less effect on the volatility than new news.

To account for these points, we have used the decay model technique that was first proposed by Yu (2014), which reflect the instantaneous impact of news releases and the decay of this impact over a subsequent period of time. The technique combines exponential decay and accumulation of the sentiment score over a given time bucket under observation. In Yu & Mitra (2016), the decay model is used to construct predictive models for return, volatility and liquidity. They modelled asset returns and asset liquidity (in terms of the bid-ask spread) using an AR(2) and AR(3) model, respectively that extended by the news impact scores. The predictive volatility model was constructed by adding the news impact scores as additional parameters (external regressors) to the variance equation of GARCH model. Further mathematical details can be found in Yu & Mitra (2016). The technique can be summarised as follows:

- Let $\{\hat{\mathbf{N}}_{\hat{t},d}^{a}\}$ be a set of all news events related to the asset a, $\hat{t}$ is time stamp and d is day, where $\hat{t} = 1, \cdots T$ and T is the last time bucket in the day. Each news item in the set is collected from the RavenPack metadata time series.

- We define a new set of filtered news events $\{\mathbf{N}_{t,d}^{a}\} \subseteq \{\hat{\mathbf{N}}_{\hat{t},d}^{a}\}$ as accepted under the filter of relevance and novelty scores arriving in the time bucket t of the trading day, where the size of a bucket is one minute and $1 \leqslant t \leqslant 510$, since the total number of minutes is 510 in a trading day.

- We map the event sentiment scores (*ESS*), which are belong to the range of [0,100], to scores that belong to the range of $[-50,50]$ for each filtered news item $\mathbf{N}_{t,d}^{a}$. We denote these mapped sentiment scores as $\mathbf{S}(\mathbf{N}_{t,d}^{a})$.

- The impact of a particular news item exponentially decays over time. We model the impact score of a news item as:

Impact Score = Event Sentiment Score X Exponential Decay Function

- For a given asset, all the relevant news items which arrived in the past have an impact on the asset price volatility at the current time bucket t. Therefore, the impact score of an old news item $(\mathbf{N}_{\tau,d}^{a})$ at current time bucket t would be calculated as:

$$\mathbf{I}_t(\mathbf{N}_{\tau,d}^{a}) = \underbrace{\mathbf{S}(\mathbf{N}_{\tau,d}^{a})}_{\text{Sentiment score}} \underbrace{e^{-\lambda(t-\tau)}}_{\text{Decay}}, \qquad t \geqslant \tau \tag{3.1}$$

where λ is the exponent which determines the decay rate. The value of λ has to be chosen such that the sentiment value decays to half the initial value in a specific time span.

- Instead of simply aggregating the impact of the news items with positive and negative impact, we keep them separate so that positive and negative effects do not cancel each other. Cancellation reduces the news flow and can lead to misinterpretation.

- We define PIS_d^a and NIS_d^a to be the sets of all the news items with positive and negative impact scores over a specific threshold for asset a on day d, respectively. In particular,

$$PIS_d^a = \left\{ \mathbf{I}_t(\mathbf{N}_{t,d}^{a}) \mid \mathbf{S}(\mathbf{N}_{t,d}^{a}) \geqslant \theta, \forall \, \mathbf{N}_{t,d}^{a} \right\} \quad \text{and} \quad NIS_d^a = \left\{ \mathbf{I}_t(\mathbf{N}_{t,d}^{a}) \mid \mathbf{S}(\mathbf{N}_{t,d}^{a}) \leqslant -\theta, \forall \, \mathbf{N}_{t,d}^{a} \right\},$$

where θ is the threshold expressed as the sentiment value that is considered large enough for inclusion in the impact computation for a given asset. In this study, we set the value of $\theta = 1$.

- To obtain positive and negative news impact scores for a particular day, d, we aggregate all positive and negative news impact scores that we have at that day separately for the asset a; such that:

$$\mathbf{P}_d^a = \sum_{t=1}^{t_m} \mathbf{I}_{t_m}(\mathbf{N}_{t,d}^a), \quad \forall \quad \mathbf{I}_{t_m}(\mathbf{N}_{t,d}^a) \in PIS_d^a,$$

and

$$\mathbf{N}_d^a = \sum_{t=1}^{t_m} \mathbf{I}_{t_m}(\mathbf{N}_{t,d}^a), \quad \forall \quad \mathbf{I}_{t_m}(\mathbf{N}_{t,d}^a) \in NIS_d^a,$$

where $t_m = 510$ is the total number of time buckets.

- We then transform these two time series for positive and negative daily impact scores into scaled news impact scores as follows. Define $\mathscr{T} = \{1, 2, \cdots, T\}$ as a finite index set for days over which data is available. Let

$$P_{(1,T)}^a = \max_{t \in \mathscr{T}} \tilde{\mathscr{P}}_t^a, \quad \text{and}$$

$$N_{(1,T)}^a = |\min_{t \in \mathscr{T}} \tilde{\mathscr{N}}_t^a|.$$

We then define

$$\mathscr{P}_t^a = \frac{\mathbf{P}_t^a}{P_{(1,T)}^a} \quad \text{and}$$

$$\mathscr{N}_t^a = \frac{\mathbf{N}_t^a}{N_{(1,T)}^a}, \tag{3.2}$$

where $t \in \mathscr{T}$. Clearly, $\mathscr{P}_t^a \in [0,1]$ and $\mathscr{N}_t^a \in [-1,0]$.

3.2 *GARCH Model*

Until early 1980s, numerous models of prediction based on autoregression were put forward. In two landmark papers Engle (1982) and Bollerslev (1986), the ARCH and GARCH (Generalized Autoregressive Conditional Heterscedasticity) models have been proposed and they are the most successful and popular models in predicting the volatility. Their incredible popularity stems from their ability to capture, with a very flexible structure, some of the typical stylized facts of financial time series, such as volatility clustering, that is the tendency for volatility periods of similar magnitude to cluster. Usually GARCH models can take into account the time-varying volatility phenomenon over a long period (see French *et al.* (1987) and Franses & Dijk (1996)) and provide very good in-sample estimates.

In an ARCH(1) model, next period's variance only depends on last period's squared residual so a crisis that caused a large residual would not have the sort of persistence that we observe after actual crises. This has led to an extension of the ARCH model to a GARCH, or Generalized ARCH model, Bollerslev (1986) and Taylor (1986) developed the GARCH(p,q) model . The model allows the conditional variance of variable to be dependent upon previous lags; first lag of the squared residual from the mean equation and present news about the volatility from the previous period which is as follows:

$$\sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^{p} \beta_i \sigma_{t-i}^2 . \tag{3.3}$$

One of the most used and the simple model is the GARCH(1,1) process, for which the conditional variance is represented as a linear function of its own lags. The simplest GARCH (1,1) model can be written as follows:

$$\text{Mean equation} \qquad r_t = \mu + \varepsilon_t, \qquad (3.4)$$

$$\text{Variance equation} \qquad \sigma_t^2 = \omega + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2, \qquad (3.5)$$

where $\omega > 0$, $\alpha_1 \geqslant 0$ and $\beta_1 \geqslant 0$. If $\alpha_1 + \beta_1 < 1$ the model is covariance stationary and the unconditional variance equals $\sigma^2 = \omega/(1 - \alpha_1 - \beta_1)$.

r_t = return of an asset at time t

μ = average return

ε_t = residual returns, defined as:

$$\varepsilon_t = \sigma_t z_t, \qquad (3.6)$$

where z_t is standardized residual returns (i.e. i.i.d random variable with zero mean and unit variance), and σ_t^2 is conditional variance.

3.3 *EGARCH Model*

In his seminal paper, Nelson (1991) pointed out that the simple structure of the GARCH and related processes, i.e. symmetric non-linear models, is not the most important criterion in modelling the financial time series. He underlined that there are some major drawbacks of the GARCH theory indicating that a new approach is necessary to be established. The Exponential GARCH (EGARCH) model of Nelson meet these limitations. This model accounts for the fact that the volatility tends to rise in response to a bad news and fall in response to a good news, by introducing a sign term in the model. Further, logarithm of conditional variance is modelled, which means that no positivity condition needs to be imposed on the parameters to guarantee positivity of conditional variance. Nelson also derived a necessary and sufficient condition for strict stationarity of the EGARCH process, when $\ln \sigma_t^2$ has an infinite moving average representation.

EGARCH(p,q) model is represented as follows:

$$\ln(\sigma_t^2) = \omega + \sum_{i=1}^{q} (\alpha_i |z_{t-i}| + \gamma_i z_{t-i}) + \sum_{j=1}^{p} \beta_j \ln(\sigma_{t-i}^2), \qquad (3.7)$$

where $z_{t-i} = \frac{\varepsilon_{t-i}}{\sigma_{t-i}}$ for $i = 1, \ldots, q$ are standardized innovations. The parameter γ_i permits asymmetric effect and if $\gamma_i = 0$ then good news ($\varepsilon_{t-i} > 0$) will have the same effect on volatility as the bad news ($\varepsilon_{t-i} < 0$). Leverage effect can be examined by testing the assumption that $\gamma_i < 0$ as negative shocks will have bigger effects on volatility than positive shocks of the same magnitude. The log function is used in EGARCH model to ensure that the process remains positive. If shocks to variance $\ln(\sigma_t^2)$ perish fast and the deterministic is removed, then $\ln(\sigma_t^2)$ is strictly stationary as shown in Nelson (1991). The formula for the simple and popular process EGARCH(1,1) is given as follows:

$$\ln(\sigma_t^2) = \omega + \alpha_1 |z_{t-1}| + \gamma_1 z_{t-1} + \beta_1 \ln(\sigma_{t-1}^2). \qquad (3.8)$$

3.4   *News Augmented GARCH Model*

In general, the basic GARCH conditional variance equation (3.3) under normality provides a reasonably good model for analyzing financial time series and estimating the conditional volatility. However, in some cases there are aspects of the model which can be improved so that it can better capture the characteristics and dynamics of a particular time series.

Before introducing a new volatility model structure to improve the volatility prediction of GARCH model using news data, we recall that trading on financial markets is strongly influenced by public company-specific, macroeconomic or political information flows. As a result, markets react sensitively to news, which is announced on both regular and irregular basis, and news events appear to affect stock return volatility quickly, suggesting that the market incorporates information quickly. The volatility tends to be higher in a falling market than in a rising market. Therefore, companies such as RavenPack and Thomson Reuters have started to automatically track and monitor relevant information on ten of thousand of companies and quantifying the content of news articles about them, and allowing for the measurement of the reaction to positive, neutral and negative news.

In our study, we are interested in developing a volatility prediction model in which enables us to use the whole information content of news events in order to improve the existing volatility prediction GARCH (1,1) model. Therefore, we are going to construct news impact scores, as derived in section (3.1), so as to be used as proxies of news events in our new model. To accomplish this, we use Raven-Pack's news analytics database and exploit its quantitative sentiment scores. Since the financial markets are mainly sensitive to good and bad news, thus we take only positive and negative news sentiments into consideration as we believe that neutral news does not have any affect on stock return volatility.

The model structure needs to reflect the following economic realities: positive and negative news impact the volatility differently. Furthermore, positive news tends to reduce volatility whereas negative news tends to increase volatility. Finally, the impact of news on return's volatility of an asset decays relatively slowly (more as a power law than as an exponential).

To model this effect of news in addition to serial correlation, we define scaled version of GARCH model, where the scaling factor is determined by the news sentiment score in the following way. Consider a function of two variables x and y:

$$f(x,y) = a + 0.5 * b \left(\frac{e^x - 1}{e^x + 1} - \frac{e^y - 1}{e^y + 1} \right), \tag{3.9}$$

where a and b are constants. It is easy to see that $f(x,y)$ lies between $(a, a+b)$ for any non-negative values of x and non-positive values of y.

Let $\{\mathscr{P}_t\}$ and $\{\mathscr{N}_t\}$ be two different time series as defined in equation (3.2)[1]. Keeping in mind the economic realities mentioned above, we define the following function as a scaling factor to GARCH model:

$$f(\mathscr{P}_t, \mathscr{N}_t) = a + 0.5 * b \left(\frac{e^{\kappa \mathscr{P}_t} - 1}{e^{\kappa \mathscr{P}_t} + 1} - \frac{e^{\gamma \mathscr{N}_t} - 1}{e^{\gamma \mathscr{N}_t} + 1} \right), \tag{3.10}$$

where 0.5 is a scaling factor of the function, and a, b, κ and γ are parameters of the model. For example, if we set the values of $a = 0.8$, $b = 0.8$, $\kappa = 4$ and $\gamma = 4$, the outcome will be as illustrated in figure 1. In this example, it can be clearly seen that the function reaches its highest value only when $\mathscr{P}_t = 1$

---

[1]The superscript for asset is omitted for simplicity since we consider only one asset at a time.
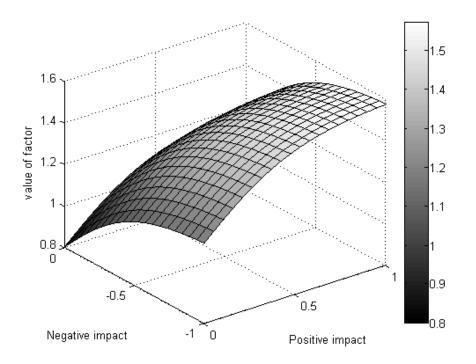
Z. SADIK, P. DATE AND G. MITRA

Figure 1. Plot of the scaling factor model against positive and negative impact scores

and $\mathcal{N}_t = -1$. This can be interpreted in our model that the volatility increases when the value of news impact scores ($\mathcal{P}_t$ and $\mathcal{N}_t$) are increasing and vice versa.

The model structure of the news augmented GARCH(1,1) model is chosen to be one with a direct multiplicative effect of news on the GARCH-predicted volatility:

$$\sigma_t^2 = \left[a + 0.5 * b \left(\frac{e^{\kappa \mathcal{P}_{t-1}} - 1}{e^{\kappa \mathcal{P}_{t-1}} + 1} - \frac{e^{\gamma \mathcal{N}_{t-1}} - 1}{e^{\gamma \mathcal{N}_{t-1}} + 1} \right) \right] (\omega + \alpha \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2), \qquad (3.11)$$

where $a > 0$, $b > 0$, $\kappa \geqslant 0$, $\gamma \geqslant 0$, $\omega > 0$, $\alpha \geqslant 0$ and $\beta \geqslant 0$. In order to keep the model covariance stationary, we have to impose this constraint $\alpha + \beta < 1$ during the time of fitting the model to the data. In addition, we also need a constraint on the upper bound of the interval range that is specified by summing up the parameters a and b to keep the news impact related scaling in a reasonable range. The choice used here is $0.5 \leqslant a + b \leqslant 2$, *i.e.* the impact of news is assumed to change the GARCH prediction at most by a factor of 2. Furthermore, ε_t is residual returns at time t and defined by:

$$\varepsilon_t = \sigma_t z_t,$$

where z_t is standardized residual returns (*i.e.* i.i.d random variable with zero mean and unit variance), and σ_t^2 is conditional variance. For the sake of brevity, from now on we call the news augmented GARCH(1,1) model as NA-GARCH model.

It is also possible to model the news impact as additive, rather than multiplicative. Our numerical experiments indicated that an additive news impact model performs a lot worse than a multiplicative news impact model. This is consistent with the intuition that news affects a percentage increase or decrease in volatility, e.g. it is conceivable that a specific negative news will cause x% increase in the *current* level of volatility, rather than causing a specific quantum of increase regardless of the current volatility.

The chosen model structure adds only four more model parameters for each asset and offers a reasonable compromise between increased model complexity and parsimony in terms of model parameters. The choice of model structure is essentially heuristic, and is justified through numerical experiments. One can also reduce the number of parameters by keeping the value of a and b fixed, for instance, $a = 0.5$ and $b = 1.5$. However, treating a and b as free parameters does improve results in terms of predictive ability of the model.

Finally, as mentioned earlier in section 1, the model structure of NA-GARCH model can easily be extended to NA-GARCH(p,q) model. We chose not to do so due to extensive evidence in the literature on the adequacy of GARCH(1,1) model for forecasting, as mentioned earlier. GARCH(1,1) model has been found to adequately fit many economic and financial time series as well as proven surprisingly successful in predicting conditional variances. Further, there is evidence in the literature that it is hard to beat GARCH(1,1) in terms of its forecasting ability and this was also confirmed in our empirical study.

4. Methodology: Calibration and Performance

In this section, we are going to explain the methods that were used in our study for parameter estimation of the models and for the analysis of the datasets of the assets. To compare the performance of the models, we will use two criteria. The first criteria is the model fit, and the second one is the prediction accuracy of the model.

4.1 *Parameter Estimation and Model Fitting*

To be able to predict the volatility for a time series, we first have to fit the model to the time series. This is done via estimation of the parameters in the our models. The most common method of this estimation has been used in our study which is the maximum-likelihood estimation (MLE). Three models have been estimated over different datasets for all assets in both indices: simple GARCH, EGARCH and NA-GARCH. In this study, the estimated parameters are initiated to the same initial assumptions for all the models. Then we estimated the parameters of the three models by using the maximum likelihood estimation, each under Gaussian distributed assumptions. To calibrate a model, maximum likelihood estimation requires us to maximize the likelihood function $L(\theta)$ with respect to the unknown parameter θ. For instance, the likelihood function of simple GARCH(1,1) model with normal conditional returns is

$$L(\alpha_0, \alpha_1, \beta_1, \mu | r_1, r_2, \cdots, r_T) = \prod_{t=1}^{T} \frac{1}{\sqrt{2\pi\sigma_t^2}} \exp\left(\frac{-(r_t - \mu)^2}{2\sigma_t^2} \right). \tag{4.1}$$

Since the logarithm is monotonically increasing the function of L, then it's equivalent to minimizing the log of the likelihood function to estimate the unknown parameters

$$\ln L(\alpha_0, \alpha_1, \beta_1, \mu | r_1, r_2, \cdots, r_T) = -\frac{T}{2} \ln(2\pi) - \frac{1}{2} \sum_{t=1}^{T} \ln \sigma_t^2 - \frac{1}{2} \sum_{t=1}^{T} \left(\frac{(r_t - \mu)^2}{\sigma_t^2} \right), \tag{4.2}$$

where σ_t^2 is given by equation (3.5) for GARCH model, the exponential of equation (3.8) for EGARCH model and equation (3.11) for NA-GARCH model.

In our study, we are using low frequency data for analysis with 750 daily returns as data points for each dataset. As mentioned in the previous sections, we are using 500 data points as in-sample period to estimate the parameters of the model, whereas the other 250 data points are kept aside as out-of-sample period for backtesting purposes. For the purpose of *in-sample* comparison of the models, we use the Akaike information criterion (AIC) to compare the goodness of fit to the data of the models. One can calculate the AIC of a model using the following formula:

$$AIC = -2ln(L) + 2P, \tag{4.3}$$

where L is the maximum value of the likelihood function for the model and P is the number of estimated parameters in the model (see Burnham & Andesron (2002), for example). While AIC tells us about in-sample performance of the model, out-of-sample performance is often far more important from a forecasting point of view. We consider the measures of out-of-sample performance in the next subsection.

4.2 *Performance Measures*

There are many statistical methods which can be used to observe the prediction accuracy of a model, for instance, Mean Absolute Error (MAE), Root Mean Square Errors (RMSE) and Mean Average Percentage Error (MAPE). In the quantitative comparison between two models in terms of prediction accuracy, the model which produces the smallest values of the forecast evaluation statistics is judged to be the best model. For this study we have considered mean absolute error and root mean square errors as the criteria for prediction accuracy since the squared daily returns may not be a proper measure to assess the forecasting performance of the different GARCH models for conditional variance (see Andersen & Bollerslev (1998)). The mean absolute error is given by

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|f_i - y_i| = \frac{1}{n}\sum_{i=1}^{n}|e_i|. \tag{4.4}$$

The root mean squared error is given by

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(f_i - y_i)^2}. \tag{4.5}$$

In both (4.4) and (4.5), f_i is the predicted value and y_i is the realized volatility, which is considered as a true value of volatility. Since the errors are squared before they are averaged, the RMSE gives a relatively high weight to large errors.

5. Computational Results

In this section, we are presenting the empirical investigation of the estimated models for all the datasets. We analyze and compare the volatility obtained by GARCH, EGARCH and NA-GARCH models described in section 3, against the observed values of volatility. We employ the realized volatility as a volatility benchmark and its values can be calculated by the standard deviation of the asset daily returns.

Typically, realised volatility is often measured as the sample standard deviation:

$$\sigma = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(r_t - \mu)^2} \tag{5.1}$$

where r_t is the return on day t and μ is the average return over the N-day period.

As we mentioned earlier in section (2.3), we have used RavenPack's news analytics database and exploit some quantitative scores of its fields in our research. In order to derive a suitable news impact time series (positive and negative impact scores) and utilize them as proxies of good and bad news, we first filtered the news metadata so as for the chosen assets, from FTSE100 and EUROSTOXX50, were selected under the filter of relevance score of 100, and any news item that had an event novelty scores under the value of 70 was ignored and not included in the dataset. Then, the event sentiment scores have been transformed into a scaled sentiment score in the range +50 to -50, as we find that such a derived single score provides a relatively better interpretation of the mood of the news item. We set the value of λ in equation (3.1) to 90 minutes, which is the exponent that determines the decay rate. The value of λ has to be chosen such that the sentiment value decays to half the initial value in a specific time span (Yu (2014)). Thus, the news sentiment score is a relative number which describes the degree of positivity and negativity in a piece of news. During the trading day, as news arrives it is given a sentiment value. Finally we followed the steps that described in section (3.1) to generate two time series of positive and negative news impact scores and use them in NA-GARCH(1,1) model in equation (3.11). Obviously, other decay functions exist. One popular choice is the Hill decay function (see Hill (1910)), which is defined as

$$f(t,\tau,\lambda) = \frac{1}{\left(1 + \frac{(t-\tau)}{\lambda}\right)^n} \qquad t \geqslant \tau \tag{5.2}$$

where t is current time, λ is the chosen value which make the sentiment value decays to half of its initial value (half-life) in a specific time span and n is the hill coefficient. We set the half-life time (λ) of 90 minutes similar to the half-life used in exponential function to generate the news impact scores. We compared the exponential decay function with the Hill decay function with the Hill coefficient n=2. We found that, in all the datasets under study, our proposed model NA-GARCH performs much better when we use the exponential function in generating the news impact scores. For instance, NA-GARCH with news impact scores that used exponential function outperforms the one with news impact scores that used Hill decay function. Tables 2 and 3 show the differences in terms of MAE and RMSE between NA-GARCH with news impact scores that used exponential function and NA-GARCH with news impact scores that used Hill function as a decay function for all chosen assets from FTSE100 and EUROSTOXX50. For example, from table 2, for the first asset, "AstraZeneca", it can be seen that our new volatility model (NA-GARCH) predicted the return volatility accurately in 9 out of 9 datasets, in terms of MAE and RMSE, using the news impact scores that were generated by using the exponential function, whereas the model predicted the return volatility accurately only in 3 out of 9 datasets using the news impact scores that were generated by using the Hill function. It is possible that using a function with a slower decay of news impact (such as the power law decay in Hill function) might be more appropriate when the news is infrequent. In our datasets, the assets are very liquid, news is frequent and correspondingly the impact of a particular news item can be assumed to die out faster than a power law. Hence the use of Hill decay function is not considered further in our study and we stick to the exponential decay function. The methodology, of course, is applicable to other decay functions.

After computing the news impact scores and calculating the returns for each asset, we calibrated the models (GARCH, EGARCH and NA-GARCH) using maximum likelihood method under the as-

	Assets	Exponential		Hill	
		AME	RMSE	AME	RMSE
1	AstraZeneca	**9**	**9**	3	3
2	Aviva	**7**	**7**	5	6
3	BP	**7**	**7**	4	3
4	GlaxoSmithKline	**8**	**9**	6	7
5	Lloyds Bank	**7**	**7**	1	2
6	Vodafone	**8**	**8**	6	5

Table 2. Contingency table for NA-GARCH model shows the performance differences of the model using different decay functions (Exponential and Hill) in terms of MAE and RMSE for the chosen FTSE100 assets. The greater of the successful cases are highlighted in boldface.

	Assets	Exponential		Hill	
		AME	RMSE	AME	RMSE
1	Allianz	**7**	**8**	5	5
2	Anheuser-Busch	**7**	**7**	6	5
3	Banco Santander	**8**	**8**	4	2
4	Bayer	**9**	**8**	4	4
5	Deutsche Bank	**8**	**8**	4	5
6	Total	**7**	**6**	4	3

Table 3. Contingency table for NA-GARCH model shows the performance differences of the model using different decay functions (Exponential and Hill) in terms of MAE and RMSE for the chosen EUROSTOXX50 assets. The greater of the successful cases are highlighted in boldface.

	Assets	MAE				RMSE			
		6	7	8	9	6	7	8	9
1	AstraZeneca	✓	✓	✓	✓	✓	✓	✓	✓
2	Aviva	✓	✓			✓	✓		
3	BP	✓	✓			✓	✓		
4	GlaxoSmithKline	✓	✓	✓		✓	✓	✓	✓
5	Lloyds Bank	✓	✓			✓	✓		
6	Vodafone	✓	✓	✓		✓	✓	✓	
Total successful cases		6	6	3	1	6	6	3	2

Table 4. Contingency table for MAE and RMSE shows the successful cases of NA-GARCH. The sub-columns represent how many times NA-GARCH model performed better than GARCH model, and the rows represent the chosen FTSE100 assets

sumption of Gaussian error distribution. We then substituted the model for σ^2 in the normal maximum likelihood and then maximized with respect to the parameters. We used the *fminsearch* function in MATLAB software to estimate the parameters.

The computational results of our empirical experiments show that the sets of coefficient estimates are quite different from one dataset to another. Tables (from 1 To 12) in the *Appendices* show the parameter estimates of the models for the returns of 12 assets from FTSE100 and EUROSTOXX50 indices. The parameters α and β for GARCH and NA-GARCH models are consistent and their sums are less than unity implying that the models are stationary, though the volatility is fairly persistent since $(\alpha + \beta)$ is close to one. In order to see the goodness of fit to the data of the models, we compared the model fit of NA-GARCH(1,1) with simple GARCH(1,1) and EGARCH(1,1) when the distribution of returns are assumed to be Gaussian.

To compare the three models in terms of goodness of fit to the data, we applied Akaike information criterion (AIC) and the Log-Likelihood value. The results show that the GARCH and EGARCH models have slightly greater log-likelihood value than NA-GARCH, which this leads the values of AIC for GARCH and EGARCH models to have slightly smaller values than NA-GARCH. In most of the cases GARCH and EGARCh have the same AIC values, and the differences between them and the NA-GARCH model in terms of AIC values were extremely small (less than 1% in average), see tables (13 - 24) in the *Appendices*. Therefore, according to the AIC values of the three models, we noticed that all of the models have almost the same level of goodness of fit.

To test the predictive ability of the models, we compared our proposed model NA-GARCH to the simple GARCH model as well as the EGARCH model for each dataset by computing MAE and RMSE, when the distribution of the returns is normal. Tables (from 13 To 24) in the *Appendices* show the MAE and RMSE results of the models for the returns of 12 assets from FTSE100 and EUROSTOXX50. In addition, if our new model NA-GARCH correctly incorporates the understanding of the impacts of news, it should be able to generate similar patterns as realized volatility in our forecasted volatility. Table 5 shows how many times our proposed model (NA-GARCH) was better than GARCH in terms of MAE and RMSE for each of the chosen assets in FTSE100. The construction of the table is as follows: there are three main columns in the table with headings "Assets", "MAE" and "RMSE". The chosen assets are listed in the first column under the heading "Assets", the second column with the heading "MAE" is further divided into four sub-columns each labelled or named with a number (from 6 to 9) that indicates the number of cases in which NA-GARCH model leads to a lower error than the GARCH model, out of the 9 datasets for each asset. For example, if the first sub-column, "6", is ticked for the first row, "AstraZeneca" asset, then this means our new volatility model (NA-GARCH) is predicted the volatility better than GARCH model in 6 out of 9 datasets in terms of MAE computations for that particular asset, and so on for the rest of sub-columns. The third column (RMSE) is also divided into four sub-columns, and has the same construction as the second column, but they are showing the number of successful cases of NA-GARCH based on the RMSE calculations. From this table, it can be seen that our new volatility model (NA-GARCH) predicted the return volatility better than GARCH model for the first asset, "AstraZeneca", in every datasets for this asset in terms of MAE and RSME; since all the related sub-columns in the first row are ticked. This means that NA-GARCH model is 100% a better choice than GARCH model to predict the volatility for this asset, whereas NA-GARCH model is predicted the volatility better than GARCH model only in 7 out of 9 datasets in terms of MAE and RMSE for the second asset, "Aviva".

We also have examined the prediction ability of the models for other six assets from EUROSTOXX50. Table 6 shows how many times our new model (NA-GARCH) was better than GARCH model in terms of MAE and RMSE for each of the chosen assets in EUROSTOXX50.

Assets		MAE				RMSE			
		6	7	8	9	6	7	8	9
1	AstraZeneca	✓	✓	✓	✓	✓	✓	✓	✓
2	Aviva	✓	✓			✓	✓		
3	BP	✓	✓			✓	✓		
4	GlaxoSmithKline	✓	✓	✓		✓	✓	✓	✓
5	Lloyds Bank	✓	✓			✓	✓		
6	Vodafone	✓	✓	✓		✓	✓	✓	
Total successful cases		6	6	3	1	6	6	3	2

Table 5. Contingency table for MAE and RMSE shows the successful cases of NA-GARCH. The sub-columns represent how many times NA-GARCH model performed better than GARCH model, and the rows represent the chosen FTSE100 Assets

Assets		MAE				RMSE			
		6	7	8	9	6	7	8	9
1	Allianz	✓	✓			✓	✓	✓	
2	Anheuser-Busch	✓	✓			✓	✓		
3	Banco Santander	✓	✓	✓		✓	✓	✓	
4	Bayer	✓	✓	✓	✓	✓	✓	✓	
5	Deutsche Bank	✓	✓	✓		✓	✓	✓	
6	Total	✓	✓			✓			
Total successful cases		6	6	3	1	6	5	4	0

Table 6. Contingency table for MAE and RMSE shows the successful cases of NA-GARCH. The sub-columns represent how many times NA-GARCH model performed better than GARCH model, and the rows represent the chosen EUROSTOXX50 Assets

Tables 7 and 8 show how many times the NA-GARCH model was able to predict the volatility better than the EGARCH model in terms of MAE and RMSE for each of the chosen assets in FTSE100 and EUROSTOXX50, respectively. All tables have the same construction. Furthermore, the results show that the NA-GARCH model correctly incorporated the understanding of the impacts of news as it was able to generate similar patterns as realized volatility in the out of sample periods for all datasets.

Assets		MAE				RMSE			
		6	7	8	9	6	7	8	9
1	AstraZeneca	✓	✓			✓	✓	✓	
2	Aviva	✓	✓	✓		✓	✓	✓	
3	BP	✓	✓			✓	✓		
4	GlaxoSmithKline	✓	✓			✓	✓	✓	
5	Lloyds Bank	✓	✓	✓		✓	✓	✓	
6	Vodafone	✓	✓			✓	✓		
Total successful cases		6	6	2	0	6	6	4	0

Table 7. Contingency table for MAE and RMSE shows the successful cases of NA-GARCH. The sub-columns represent how many times NA-GARCH model performed better than EGARCH model, and the rows represent the chosen FTSE100 Assets

Assets		MAE				RMSE			
		6	7	8	9	6	7	8	9
1	Allianz	✓	✓	✓		✓	✓		
2	Anheuser-Busch	✓	✓			✓	✓		
3	Banco Santander	✓	✓	✓		✓	✓	✓	
4	Bayer	✓	✓	✓	✓	✓	✓	✓	✓
5	Deutsche Bank	✓	✓	✓		✓	✓	✓	
6	Total	✓	✓	✓		✓	✓	✓	
Total successful cases		6	6	5	1	6	6	4	1

Table 8. Contingency table for MAE and RMSE shows the successful cases of NA-GARCH. The sub-columns represent how many times NA-GARCH model performed better than EGARCH model, and the rows represent the chosen EUROSTOXX50 Assets

To sum up, from the tables 5 and 6 we can clearly see that NA-GARCH performs better than GARCH model in terms of prediction accuracy in at least two thirds (6 out of 9) or more of the datasets for each asset. Overall, there are 216 comparisons, with 9 datasets, 12 assets, 2 error metrics. Out of these, NA-GARCH model turns out to be a better model than the GARCH model in 184 comparisons. Further, it is obvious from the tables 7 and 8 that NA-GARCH outperforms the EGARCH model in terms of prediction accuracy in at least two thirds (7 out of 9) or more of the datasets for each asset. Again, out of 216 comparisons, NA-GARCH model turns out to be a better model than the EGARCH model in 185 comparisons. Furthermore, all the sample means were positive, for instance, the sample means of differences ($MAE_{GARCH} - MAE_{NA-GARCH}$) and ($RMSE_{GARCH} - RMSE_{NA-GARCH}$) are both positive, for both indices FTSE100 and EUROSTOXX50. This analysis suggests that including the news impact term in the GARCH framework has improved the predictive ability of GARCH model.

6. Discussion and Conclusion

Forecasting accurately future volatility and correlations of financial asset returns is essential since volatility can be used on its own such as in the hedge fund portfolio, or it can be used in conjunction with return measures such as in Sharpe and Sortino ratio formulae. Therefore, the flexibility of GARCH model and its forecasting accuracy, place the model in a unique position to achieve many of the requirements of the practitioners. However, its use is restricted to one time series data (market data) and ignoring other types of data, especially the market news sentiment that can be very helpful to anticipate the potential impacts on the return volatility of an asset.

This paper investigates the impact of high-frequency public news sentiment on the daily log returns volatility for twelve assets from FTSE100 and EUROSTOXX50; a time period from 3 January 2005 to 31 December 2015. We propose a new volatility model, namely News Augmented GARCH (NA-GARCH) model, which enables us to use the news sentiment score to improve the predictive ability of GARCH model. We analysed the computational results of the empirical investigation of the two sets of six assets comparing the performances of the GARCH, EGARCH and NA-GARCH models using the chosen performance measures. Our results clearly demonstrate that NA-GARCH provides a superior prediction of volatility than the simple GARCH and EGARCH models. Our empirical results suggest that, when news sentiment score is available, NA-GARCH is a far better choice than GARCH and EGARCH for volatility prediction. In particular, the NA-GARCH model incorporates the primary characteristics of historical return volatility clustering with news sentiment scores, which make it capable of capturing the more general features of volatility and provide more robust and transparent predictive abilities over longer, out of sample time horizons. The study also shows that the NA-GARCH model correctly incorporated the available information concerning news impact into our forecasts. The patterns that generated by NA-GARCH presented the same as realised volatility. We have used pre-processed news sentiment data from commercial provider *RavenPack*. It is reasonable to expect that a volatility prediction model and methodology developed in our paper, which works with news sentiment score from one data provider, would also work with score from another data provider. Our findings also show that the positive news tends to reduce volatility whereas negative news tends to increase volatility, which is consistent with the studies by Crouhy & Rockinger (1997), Song (2010) and Chen & Ghysels (2011) about the effect of positive and negative shocks. This analysis suggests that including the news impact term in the GARCH framework has improved the predictive ability of GARCH model. Another suggestion was the use of exponential decay function is good when the news flow is frequent, whereas the Hill decay function is good only when there are scheduled announcements, because the impact of the news on the market dies slowly as the impact of new information on a stock market depends on how unexpected the news is. NA-GARCH is thus a computationally efficient means of exploiting the news sentiment score for better volatility prediction and it has a potential to be very useful in industrial practice. Therefore, our findings are crucial for all investors who are trading on the variance or volatility swaps, which can be used to speculate on future realized volatility, or to hedge the volatility exposure of other positions since the profit and loss from a variance swap depends directly on the difference between realized and implied volatility of a given underlying asset. In addition, NA-GARCH model would be useful for investors who are focusing on risk-adjusted returns, especially those that utilize asset allocation and volatility targeting strategies. Furthermore, NA-GARCH model can be used to estimate Value-at-Risk more accurately. An empirical study of VaR prediction and a Bayesian update strategy for coefficients of the NA-GARCH model are topics of current research.

REFERENCES

ANDERSEN, T. & BOLLERSLEV, T. (1998) Answering the skeptics: yes, standard volatility models do provide accurate forecasts, *International Economic Review,* **39**, 885-905.

ARBEX-VALLE, C.; ERLWEIN-SAYER, C.; KOCHENDRFER, A.; KBLER, B.; MITRA, G.; NANA, G.; NOUWT, B. & STALKNECHT, B. (2013) News-Enhanced Market Risk Management. Available at SSRN: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2322668.

ARSHANAPALLI, B., d'OUVILLE, E.; FABOZZI, F. & SWITZER, L. (2006) Macroeconomic news effects on conditional volatilities in the bond and stock markets, *Applied Financial Economics,* **16**(5), 377–384.

BAILLIE, R. & BOLLERSLEV, T. (1989) The message in daily exchange rates: a conditional variance tale. *Journal of Business Economic Statistics*, **7**, 297–305.

BOLLERSLEV, T. (1986) Generalized Autoregressive conditional heteroscedasticity, *The Review of Economics and Statistics,* **69**(3), 542–547.

BOLLERSLEV, T. (1987) A Conditionally Hetroskedastic Time Series Model for Speculative Prices and Rates of Return, *Journal of Econometrics,* **31**, 307–327.

BURNHAM, K.P. & ANDERSON, D.R. (2002) *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach (2nd ed.),* Springer-Verlag, ISBN 0-387-95364-7.

CHEN, X. & GHYSELS, E. (2011) NewsGood or Badand Its Impact on Volatility Predictions over Multiple Horizons, *The Review of Financial Studies,* **24**(1), 46-81.

CROUHY, M. & ROCKINGER, M. (1997) Volatility Clustering, Asymmetry and Hysteresis in Stock Returns: International Evidence, *Financial Engineering and the Japanese Markets,* **4**, 1–35.

ENGLE, R.F. (1982) Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica ,* **50**(4), 987–1007.

ENGLE, R.,& NG, V. (1993). Measuring and Testing the Impact of News on Volatility. *The Journal of Finance*, **48**(5), 1749–1778.

FENG, Y., BERAN, J. & YU, K. (2007) Modelling financial time series with SEMIFARGARCH model, *IMA Journal of Management Mathematics*, **18**, 395–412.

FRANSES, P.H., & DIJK, R.V. (1996) Forecasting Stock Market Volatility Using (Non-Linear) GARCH Models, *Journal of Forecasting,* **15**(3), 229–235.

FRENCH, K.R., SCWERT, G.W. & STAMBAUGH, R. F. (1987) Expected Stock Returns and Volatility, *Journal of Financial Economics,* **19**, 3–30.

GLOSTEN, L.R., JAGANNATHAN, R. & RUNKLE, D. (1993) On the Relation Between the Expected Value and the Volatility of the Nominal Excess Return on Stocks, *Journal of Finance,* **48**, 1779–1801.

GRAY, S.F. (1996) Modeling the Conditional Distribution of Interest Rates as a Regime-Switching Process, *Journal of Financial Economics,* **42**, 27–62.

HANSEN, P. R., & LUNDE, A. (2005) A forecast comparison of volatility models: does anything beat a GARCH (1, 1)?, *Journal of applied econometrics,* **20**(7), 873–889.

HILL, A. V. (1910) The possible effects of the aggregation of the molecules of hmoglobin on its dissociation curves, *Journal of Physiology,* **40** iv–vii.

HO, K.Y., SHI, Y., & ZHANG, Z. (2013) How does news sentiment impact asset volatility? Evidence from long memory and regime-switching approaches, *The North American Journal of Economics and Finance,* **26**, 436–456.

HSIEH, D. (1988) The statistical properties of daily foreign exchange rates: 1974-1984, *Journal of International*

Economics, **24**, 129–45.

HSIEH, D. (1989) Modelling heteroskedasticity in daily foreign exchange rates. *Journal of Business Economic Statistics*, **1**, 307–17.

KALEV, P.S., LIU, W.M.; PHAM, P.K.; & JARNECIC, E. (2004) Public information arrival and volatility of intraday stock returns, *Journal of Banking & Finance,* **28**(6), 1441–1467.

LI, LI & ENGLE, ROBERT F. (1998) *Macroeconomic Announcements and Volatility of Treasury Futures*. Department of Economics, UCSD. UC San Diego: Department of Economics, UCSD. Retrieved from: http://escholarship.org/uc/item/7rd4g3bk.

MCCURDY, T. & MORGAN, I. (1987) Tests of Martingale hypothesis for foreign currency futures with time varying volatility. *International Journal of Forecasting*, **3**, 131–48.

MITCHELL, M.L., & MULHERIN, J.H. (1994) The impact of public information on the stock market, *Journal of finance,* **49**(3), 923–950.

MITRA, L., MITRA, G. & DIBARTOLOMEO, D. (2009) Equity portfolio risk (volatility) estimation using market information and sentiment, *Quantitative Finance,* **9**(8), 887–895.

MITRA, G. & MITRA, L. (2011) *The Handbook of News Analytics in Finance,* Wiley Finance.

NELSON, D.B. (1991) Conditional heteroskedasticity in asset returns: A new approach, *Journal of the Econometric Society,* **59**(2), 347–370.

RAVENPACK (2014) *RavenPack News Analytics, Version 4.0: User Guide and Service Overview.*

RIORDAN, R., STORKENMAIER, A.; WAGENER, M. & ZHANG, S. S. (2013) Public information arrival: Price discovery and liquidity in electronic limit order markets, *Journal of Banking & Finance,* **37**(4), 1148–1159.

SHARMA, J.L., MOUGOUE, M. & KAMATH, R. (1996), Heteroscedasticity in stock market indicator return data: volume versus GARCH effects, *Applied Financial Economics*, **6**(4), 337–342.

SIDOROV, S., DATE, P. & BALASH, V. (2013) Using news analytics data in GARCH models, *Applied Econometrics*,**29**, 82–96.

SIDOROV, S., DATE, P. & BALASH, V. (2014) GARCH Type Volatility Models Augmented with News Intensity Data, *Chaos, Complexity and Leadership*, Springer Netherlands, 199–207.

SONG, JOON-YOUNG(2010) Asymmetry impact of news on stock return volatility, *Banking and Capital Markets, New International Perspectives*, World Scientific, 373–410.

TETLOCK P.C. (2007) Giving content to investor sentiment: The role of media in the stock market, *Journal of Finance,* **62**(3), 11391168.

TETLOCK, P.C. (2010) Does public financial news resolve asymmetric information?, *Review of Financial Studies,* **23**(9), 3520–3557.

THOMSON REUTERS (2010) Thomson Reuters News Analytics. Version 2.0.2.

TSAY, R.S. (2008) *Out-Of-Sample Forecasts.* Lecture Notes, The University of Chicago Booth School of Business. http://faculty.chicagobooth.edu/ruey.tsay/teaching/bs41202/sp2008/lec4a-08.pdf. Accessed January 2017.

YU, X. (2014) Analysis of news sentiment and its applications to finance. *Ph.D. Thesis,* Brunel University London, UK.

YU, X. & MITRA, G. (2016) An Impact Measure for News: its use in (daily) trading strategies, *The Handbook of Sentiment Analysis in Finance.* pp. 288–309.

ZAKOIAN, J.M. (1994) Threshold heteroskedastic models, *Journal of Economic Dynamics and Control,* **18**(5), 931–955.