# News Sentiment And States of Stock Return Volatility: Evidence from Long Memory and Discrete Choice Models

# Y. Shi a, K.Y-. Ho a

<sup>a</sup>Research School of Finance, Actuarial Studies and Applied Statistics, The Australian National University, Canberra, Australia Email: <u>yanlin.shi@anu.edu.au</u>

**Abstract:** This paper examines the relationship between the states of firm-level return volatility and public news sentiment. To incorporate structural breaks into modelling the long memory property of stock return volatility, we develop FIGARCH models that allow the constant in the conditional variance to vary with time. Following this train of thought, we firstly propose a Markov Regime-Switching FIGARCH (MRS-FIGARCH) model that allows the constant term to switch between low and high volatility states. This model is shown to outperform the Adaptive FIGARCH and Time-Varying FIGARCH models with respect to model comparison criteria. It is subsequently used to estimate the smoothing probability and the conditional variance. Second, states of firm-level return volatility are identified by comparing the previously generated smoothing probability with certain thresholds. Then, we employ discrete choice models to investigate the impact of public news sentiment on the volatility states of hourly returns of constituent stocks in the Dow Jones Composite Average (DJN 65). Our news dataset is constructed from the new RavenPack Dow Jones News Analytics database that captures over 1200 types of firm-specific and macroeconomic news releases and their sentiment scores at high frequencies. Estimated results show that news sentiment has various significant effects on the likelihood of volatility states of intraday stock return.

Keywords: Public Information Arrival, Asset Volatility, News Sentiment, Markov Regime-Switching FIGARCH, Discrete Choice Model

# **1 INTRODUCTION**

The relationship between public information arrival/news flow and asset volatility has been a focus of research on finance in the past few decades. Most of the studies are theoretically motivated by the "Mixture of Distributions Hypothesis" (MDH), which provides an explanation for the relationship between the rate of information arrival and measures of market activity, such as asset volatility and trading volume<sup>1</sup>. Among them, a recent study Ho et al. (2013) extends this topic to the impact of news sentiment on asset volatility. They argue that both positive and negative firm-specific and macroeconomic news can significantly and positively affect the asset volatility with asymmetrical effects. Furthermore, the effects are different in low-volatility and high-volatility states.

Based on and extended from their study, this paper aims to address the following set of research questions. How can the flow and sentiment of the high-frequency announcements account for the likelihood of higher volatility state of intraday asset? Are there any differences between the impact of negative news sentiment and positive news sentiment on volatility? Is the impact of firm-specific announcements stronger or weaker than that of macroeconomic news?

We study the impact of newswire messages on volatility states of stocks traded on the Dow Jones Composite Average. These newswire messages are obtained from the RavenPack News Analytics - Dow Jones Edition, which is a comprehensive database covering more than 1200 types of firm-specific and macroeconomic news events<sup>2</sup>.

To analyze the relationship between news flow and volatility states of stocks, we need to model how volatility evolves. In Ho et al. (2013), they employ two different frameworks, the FIGARCH model proposed by Baillie et al. (1996) and the Markov Regime-Switching GARCH (MRS-GARCH) model, which is an extension of GARCH model from the idea of Hamilton (1989). Also, they argue that the FIGARCH framework generally outperforms the MRS-GARCH framework in terms of the model performance. However, Diebold and Inoue (2001) argue that when structural change or stochastic regime switching exist, they are related to long memory and are easily confused with it. In addition, Ho et al. (2013) also point out that structural change widely exists in real economy. Therefore, structural change should be incorporated into the FIGARCH model to avoid spurious regression.

In order to model the structural change in FIGARCH framework, we firstly employ Adaptive FIGARCH (A-FIGARCH, proposed by Baillie and Morana (2009)) model and Time-Varying FIGARCH (TV-FIGARCH, proposed by Belkhouja and Boutahary (2011)) in our research. Both of them allow the intercept in the conditional equation to be time-varying. Furthermore, we adopt the seminal idea of Hamilton (1989) to allow regime-switching of intercept in the conditional equation of FIGARCH model and propose the Markov Regime-Switching FIGARCH (MRS-FIGARCH) model. All the three FIGARCH family models are then fitted and compared. It is demonstrated that the MRS-FIGARCH model outperforms the others in terms of model performance. Therefore, MRS-FIGARCH model is used to estimate the stock return volatility and identify the low- and high-volatility states.

To proxy news sentiment, we adopt the same approach in Ho et al. (2013) to construct the weighted number of news variables, where the weight is decided by relevance and sentiment of news. And news variables are composed of firm-specific and macroeconomic news. Discrete choice models are then employed to model the impact of news variables on volatility states. Volatility states of stock return are identified by the smoothing probability estimated from MRS-FIGARCH model. They are further classified into "calm" (low-volatility), "mediate" (mediate-volatility) and "turbulent" (high-volatility) states. Firstly, our data are fitted by dummy news variables (indicating negative or positive news) into ordered logit and probit models. The results show that occurrences of news can significantly affect the likelihood of higher volatility states. Then, we constrain our sample to periods during which at least one positive or negative news story is received. In the constrained sample, states are classified in the same way into three states and are further fitted by the weighted number of news variables.

Nonparametric tests including Wilcoxon signed rank test and Bootstrap percentile confidence interval are used to test the marginal effects of news. We show that both number and sentiment score of news can further significantly affect the likelihood of higher volatility states. Generally, after considering the sentiment and number of news, macroeconomic news has similar marginal effect as firm-specific news. More specifically,

<sup>&</sup>lt;sup>1</sup>See Ho et al. (2013) for a complete literature review of this topic.

<sup>&</sup>lt;sup>2</sup>See Ho et al. (2013) for details of RavenPack.

negative news increase the likelihood of higher volatility states, whereas positive news decreases that to a larger degree. To robust our results, we reclassify the states into two cases: "calm" and "turbulent" states based on smoothing probability. Binary logit and probit models are then employed to fit the data. We demonstrate that results are consistent across different classifications and models.

Our results suggest that both negative and positive macroeconomic news can significantly affect the volatility states of stock return volatility. Since macroeconomic news contains sentiment information of real-time macroeconomic state, our results also imply that macro economy is the key to affect the volatility state of individual stock. The results can provide useful information for professional traders who read newswires like Dow Jones. They spend a considerable amount of money on these information sources and emphasize the importance of speed and accuracy of news. Newswire messages represent much of the overall information traders receive on a real-time basis. The intraday impact of these newswire messages will help traders anticipate the potential effects on volatility states of the assets that they are monitoring (Ho et al., 2013).

#### **2** DATA AND SAMPLE

#### 2.1 Return Series

Our stock price series comprise the hourly data of Dow Jones 65 stocks<sup>3</sup> (DJN 65) from January 1, 2000 to December 31, 2011. These series are sourced from the Thomson Reuters Tick History (TRTH) database, which contains high-frequency stock data and is provided by the Securities Industry Research Centre of Australasia (SIRCA). Define  $\{S_0, S_1, \ldots, S_{T-1}, S_T\}$  as the sequence of hourly closing price for one firm at times  $\{0, 1, \ldots, T-1, T\}$ . The corresponding hourly return at time t is

$$r_t = 100 \times \log(S_t / S_{t-1}) \tag{1}$$

#### 2.2 Public Information Arrival

To proxy for public information arrivals, we use the RavenPack News Analytics database, which captures more than 1200 types of intraday firm-specific and macroeconomic news events starting from January 2000 onwards. Among dozens of the analytical outputs in the database, we use the news relevance score (*REL*), composite sentiment score (*CSS*), event sentiment score (*ESS*) and event novelty score (*ENS*)<sup>4</sup>

We construct four weighted number of news variables as done in Ho et al. (2013) using *REL*, *ESS*, *ENS* and *CSS*. At time *t*, the weighted numbers of negative and positive firm-specific and negative and positive macro-economic news stories are:

$$WNN_{f,t} = \sum_{all \ \tau} \frac{I(CSS_{f,\tau} < 50) |CSS_{f,\tau} - 50| REL_{f,\tau}}{100}$$
(2)

$$WNN_{m,t} = \sum_{all \ \tau} \frac{I(ESS_{m,\tau} < 50) |ESS_{m,\tau} - 50| ENS_{m,\tau}}{100}$$
(3)

$$WNP_{f,t} = \sum_{all \ \tau} \frac{I(CSS_{f,\tau} > 50) |CSS_{f,\tau} - 50| REL_{f,\tau}}{100}$$
(4)

$$WNP_{m,t} = \sum_{all \ \tau} \frac{I(ESS_{m,\tau} > 50) |ESS_{m,\tau} - 50| ENS_{m,\tau}}{100}$$
(5)

where subscripts f and m indicate that the corresponding scores (variables) refer to firm-specific and macroeconomic news, respectively. WNN is the negative news and WNP is the positive news.  $I(\cdot)$  is the indicator function that gives 1 when the condition inside the brackets is true and gives 0 otherwise,  $\tau$  is the time at which the news story is received and  $\tau \in [t - 1, t]$ . Furthermore, we construct four dummy news variables. Let the dummy variable be 1 when the corresponding weighted number of news variable is not 0, and set it to 0 otherwise. In this way, we use  $DN_{f,t}$ ,  $DP_{f,t}$ ,  $DN_{m,t}$  and  $DP_{m,t}$  to indicate dummy variables of the negative and positive firm-specific and negative and positive macro-economic news, respectively.

 $<sup>^{3}</sup>$ To ensure that there are sufficient observations and the datasets are complete over the sample period, 50 out of 65 firms are included in our final sample. Our selection criteria are: (1), in each month, there are at least 15 trading days; and (2), there must be no missing month in the whole period. These criteria are also employed by Ho et al. (2013).

<sup>&</sup>lt;sup>4</sup>See Ho et al. (2013) for detailed definition of those scores. The algorithms of how those scores are generated can be found in their appendix.

## **3** METHODOLOGY AND MODEL SPECIFICATION

#### 3.1 The original FIGARCH Model

FIGARCH model is proposed by Baillie et al. (1996), which is extended from GARCH family models to model the time-varying volatility. The original FIGARCH(1,d,1) model is described as follows.

$$r_{t} = \mu + \varepsilon_{t}, \, \varepsilon_{t} = \eta_{t} \sqrt{h_{t}} \text{ and } (1 - b_{1}L)h_{t} = \omega + \left[ (1 - b_{1}L) - (1 - \varphi_{1}L)(1 - L)^{d} \right] \varepsilon_{t}^{2}$$
(6)

where  $\varepsilon_t$  is the error at time t.  $h_t$  is the conditional volatility of  $\varepsilon_t$  at time t.  $\eta_t$  is an identical and independent sequence following a specific distribution. L is the lag operator.  $(1-L)^d$  is the fractional differencing operator. d is the long memory parameter. We have a stationary long memory process when 0 < d < 1. If d = 1, the process has a unit root and thus a permanent shock effect, which is equivalent as the IGARCH model. If d = 0, the process reduces to an ordinary GARCH process without long memory property (Baillie et al., 1996).

### 3.2 A-FIGARCH Model

The main weakness of the original FIGARCH model is that it assumes that the conditional volatility has only one regime over the entire period, which can lead to spurious regression. In order to overcome such drawback, Baillie and Morana (2009) points out that a powerful approach is to allow the intercept to be time dependent. Compared with the FIGARCH model, only the conditional equation is changed, which is described as follows.

$$b(L)h_t = \omega + \left[b(L) - \varphi(L)(1-L)^d\right]\varepsilon_t^2 + \omega_t \text{ and } \omega_t = \sum_{j=1}^k \left[\gamma_j \sin(2\pi jt/T) + \delta_j \cos(2\pi jt/T)\right]$$
(7)

where T is the number of observations. Moreover, they argue that adequate approximations can be achieved with very parsimonious specifications of only k = 1 or 2.

## 3.3 TV-FIGARCH Model

Belkhouja and Boutahary (2011) provide another approach to model the variation in the intercept and propose the TV-FIGARCH model. The conditional equation of their model is constructed as follows.

$$b(L)h_t = \omega + \left[b(L) - \varphi(L)(1-L)^d\right] \varepsilon_t^2 + \sum_{r=1}^R \omega_r F_r(s_t, \gamma_r, c_r)$$
  
where  $s_t = t/T$  and  $F_r(s_t, \gamma_r, c_r) = (1 + \exp\{-\gamma_r(s_t - c_r)\})^{-1}$  (8)

In addition,  $\gamma_r$  controls the degree of smoothness and must be positive, while  $c_r$  is the threshold parameter with constrain  $c_1 \leq c_2 \leq \ldots \leq c_R$ .  $s_t = t/T$  is the transition variable. When  $\gamma_r \to \infty$ , the switch from one state to another is abrupt, that is, a smooth change approaches a structural break at the threshold parameter  $c_r$ .

#### 3.4 MRS-FIGARCH Model

Apart from allowing intercept to be time dependent, another possible pathway is the idea of regime-switching, which is firstly proposed Hamilton (1989) to allow transitions between state spaces. In order to retain the advantages of GARCH models and allow the structure breaks at the same time, MRS-GARCH model is developed from Hamilton (1989)'s work<sup>5</sup>.

In this paper, we expand the two-state MRS-GARCH model used in Ho et al. (2013) to FIGARCH framework and allow the intercept term in the conditional variance equation to switch between states. The proposed MRS-FIGARCH model with Student-t innovation is as follows.

$$r_{t} = \mu + \varepsilon_{t}$$

$$\varepsilon_{t} = \eta_{t} \sqrt{h_{s_{t},t}} \text{ where } \eta_{t} \stackrel{iid}{\sim} t(0,1,v)$$

$$b(L)h_{s_{t},t} = \begin{cases} \omega_{1} + \begin{bmatrix} b(L) - \varphi(L)(1-L)^{d} \\ \omega_{2} + \begin{bmatrix} b(L) - \varphi(L)(1-L)^{d} \end{bmatrix} \varepsilon_{t}^{2} \text{ when } s_{t} = 1$$

$$\varepsilon_{t}^{2} \text{ when } s_{t} = 2 \end{cases}$$
(9)

where  $s_t$  is the state the stock lies in at time t.  $h_{s_t,t}$  is the conditional volatility in state  $s_t$  at time t. v is

<sup>&</sup>lt;sup>5</sup>See Ho et al. (2013) for a summary of the development of MRS-GARCH model.

the degree of freedom of the Student-t distribution<sup>6</sup> Also, the sequence  $\{s_t\}$  is assumed to be a stationary, irreducible Markov process with discrete state space  $\{1,2\}$  and transition matrix  $P = [p_{jk}]$  where  $p_{jk} = P(s_{t+1} = k | s_t = j)$  is the transition probability of moving from state j to state k ( $j, k \in \{1,2\}$ ). Due to the reason that unconditional variance of FIGARCH model does not exist, we constrain  $\omega_1 < \omega_2$ . Then, state 1 is the "calm" state (low-volatility state) and state 2 is the "turbulent" state (high-volatility state) in our study. The parameters of the MRS-FIGARCH model are estimated using the maximum likelihood estimation (MLE) described in Ho et al. (2013).

In order to identify which economic state the stock return lies in at time t, we extract the smoothing probability of "calm" state as follows (Hamilton, 1989).

$$P(s_t = 1|\theta, \Omega_T) = \omega_{1,t} \left[ \frac{p_{11}P(s_{t+1} = 1|\theta, \Omega_T)}{P(s_{t+1} = 1|\theta, \Omega_t)} + \frac{p_{12}P(s_{t+1} = 2|\theta, \Omega_T)}{P(s_{t+1} = 2|\theta, \Omega_t)} \right],$$
(10)

where  $\Omega_{t-1}$  is the information set at time t-1.  $\theta$  is the vector of parameters.  $\omega_{j,t-1}$  is the filtered probability in state j at time t-1, which is equal to  $P(s_{t-1} = j | \theta, \Omega_{t-1})^7$ . Using the fact that  $P(s_T = 1 | \theta, \Omega_T) = \omega_{1,T}$ , the smoothing probability series  $P(s_t = 1 | \theta, \Omega_T)$  can be generated by iterating (10) backward from T to 1.

#### 4 EMPIRICAL RESULTS AND IMPLICATIONS

#### 4.1 The FIGARCH Family Models

The conditional volatility  $h_t$  from FIGARCH family models is used to proxy the latent volatility of stock return at time t. Therefore,  $h_t$  is firstly estimated respectively from A-FIGARCH, TV-FIGARCH and MRS-FIGARCH model. We will choose the estimated  $h_t$  from the best model according to the model performance evaluation.

Table 1. Mean Estimates of Long Memory Parameter and Model Performance for FIGARCH Family Models

	d	MSE	AIC	BIC
A-FIGARCH	0.4917	14.8560	37711	37775
TV-FIGARCH	0.4914	14.8464	37699	37770
MRS-FIGARCH	0.3996	13.9112	37672	37744

Long memory parameter d is the most important parameter in FIGARCH family models. Therefore, the estimates of d from three FIGARCH family models are summarized in Table 1<sup>8</sup>. According to Diebold and Inoue (2001), if the structural change can be correctly modelled, the estimate of d is expected to be smaller. In addition, we compare the forecasting performances of the three models. The in-sample forecasting performances are compared by firstly constructing the intra-period high-low range volatility (*RNV*)  $RNV_t = [\log(\max\{S_{\tau}\} - \min\{S_{\tau}\}) \times 100]^2/(4/\log 2)$  where  $\{S_{\tau}\}$  is the sequence of intra-period stock prices between the hours t - 1 and t. Then, the mean squared error (*MSE*) of the forecast performances are defined as  $MSE = [\sum_{t=1}^{T} (RZV_t - h_t)^2]/T$  where  $h_t$  is the estimated conditional volatility<sup>9</sup> from FIGARCH family models at time t and T is the sample size. The *MSE*, along with the widely used model selection statistics Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) are generated for all the three FIGARCH family models. The results are summarized in Table 1<sup>10</sup>.

From Table 1, MRS-FIGARCH model leads to the smallest mean for estimate of d and MSE, AIC and BIC. As a result, MRS-FIGARCH model generally outperforms both A-FIGARCH and TV-FIGARCH models

<sup>&</sup>lt;sup>6</sup>The reason to use Student-t distribution is that in real economy, financial series is rarely Gaussian but typically leptokurtic and exhibits heavy-tail behaviour. Therefore, distributions which can capture those properties will lead to more efficient estimation. Also, it is argued that if regimes are not Gaussian but leptokurtic, the use of within-regime normality can seriously affect the identification of the regime process. See footnote 9 in Ho et al. (2013) for relevant references. Therefore, we use Studnet-t distribution instead of the Gaussian distribution.

<sup>&</sup>lt;sup>7</sup>See Hamilton (1989) for how to estimate the filtered probability.

<sup>&</sup>lt;sup>8</sup>To make the models parsimonious, we set both K and R to be 1 for A-FIGARCH and TV-FIGARCH models, respectively, in all cases.

<sup>&</sup>lt;sup>9</sup>For MRS-FIGARCH model, estimated volatility is the weighted average of estimated volatility in "calm" and "turbulent" states. The weights are set to the corresponding update probabilities  $P(s_t|\theta, \Omega_{t-1})$ . The details can be found in Hamilton (1989).

<sup>&</sup>lt;sup>10</sup>Only the mean is reported here, and other statistics are available upon request

with respect to the modelling of structural change and the evaluation of model performance. Hence, We are going to use  $h_t$  estimated from MRS-FIGARCH model to proxy the latent volatility of stock return. Also,  $P(s_t = 1|\theta, \Omega_T)$  is used to identify states of the stock return volatility.

# 4.2 The Effect of News on Stock Return Volatility<sup>11</sup>

In this section, we will test the effect of news on volatility states of stock return via the discrete choice model.

**Dummy News variables on Full Dataset.** To identify the states of stock return, Wilfling (2009) suggests that the following rule should be applied to the smoothing probability  $P(s_t = 1 | \theta, \Omega_T)$ : stock return volatility lies in extremely low-volatility state when  $P(s_t = 1 | \theta, \Omega_T)$  is close to 1, lies in extremely high-volatility state when  $P(s_t = 1 | \theta, \Omega_T)$  is close to 0 and lies in intermediate state otherwise. Therefore, let  $Y_t$  be the discrete variable indicating the state at time t, we set  $Y_t = 1$  indicating "calm" state when  $P(s_t = 1 | \theta, \Omega_T) \ge 0.8$ , set  $Y_t = 3$  indicating "turbulent" state when  $P(s_t = 1 | \theta, \Omega_T) \le 0.2$  and set  $Y_t = 2$  indicating "mediate" state otherwise<sup>12</sup>.

The following two equations are fitted respectively with ordered logit and probit models as follows <sup>13</sup>:

Null model: 
$$U_t = \beta_0 + \varepsilon_t$$
  
Full model:  $U_t = \beta_0 + \beta N e w s_t + \varepsilon_t$  (11)

where  $U_t$  is the latent utility of  $Y_t$ , the form of which will be decided by the logit and probit models.  $News_t$ is the news variable vector, including  $DN_{f,t}$ ,  $DP_{f,t}$ ,  $DN_{m,t}$  and  $DP_{m,t}$  in this case. The mean of estimated coefficients of dummy news variables are reported in Table 2. In order to test their statistical significance, we employ the Wilcoxon signed rank test and 95% Bootstrap percentile confidence interval<sup>14</sup>.<sup>15</sup> From Table 2, it is clear that means are mostly significantly different from 0 at 5% level (except for  $DP_{f,t}$ ), confirmed by both Wilcoxon test and Bootstrap confidence interval. Moreover, the occurrences of negative firm-specific and positive macroeconomic news have significantly negative effects, whereas negative macroeconomic news has significantly positive effect. Furthermore, marginal effects of two macroeconomic news variables are roughly close and greater than that of  $DN_{f,t}$ . Therefore, we conclude that the occurrences of news can significantly affect the likelihood of volatility states of stock return, and macroeconomic news has greater effect than firmspecific news.

Table 2.	Nonparametric	Tests of Dumm	y and Weighted	l Number of News	Variables in	Ordered Logit Model
			/ //			

Var.	Mean	Boot-C.I.	Var.	Mean	Boot-C.I.
$DN_{f,t}$	-0.1867*	[-0.2208 , -0.1001]*	$WNN_{f,t}$	0.0030*	[0.0022, 0.0058]*
$DP_{f,t}$	-0.0370	[-0.0615, 0.0336]	$WNP_{f,t}$	-0.0036*	[-0.0050, 0.0006]
$DN_{m,t}$	0.2119*	[0.2019, 0.2415]*	$WNN_{m,t}$	0.0027*	[0.0025, 0.0030]*
$DP_{m,t}$	-0.2181*	[-0.2277 , -0.1890]*	$WNP_{m,t}$	-0.0036*	[-0.0038 , -0.0032]*

Weighted Number of News variables on Constrained Dataset. In last section, we have demonstrated that occurrences of different news have various significant effects on volatility states of stock return. As described in Ho et al. (2013), it is argued that apart from the occurrence of news, number and sentiment of news stories can further affect stock return volatility.

In order to test such argument, we firstly constrain our dataset from entire period to the period containing at least one negative or positive news story is received. In other words, we constrain our sample to be the time *i*, where  $i \in \{0, 1, 2, ..., T\}$  and  $I(DN_{f,i} + DP_{f,i} + DN_{m,i} + DP_{m,i} > 0) = 1, \forall i$ .

<sup>&</sup>lt;sup>11</sup>The results presented in this section are consistent with those from two-state classification with binary logit and probit models, where threshold probability is 0.5. Estimates are available upon request.

<sup>&</sup>lt;sup>12</sup>We also perform our analysis on the choices of other thresholds, such as 0.75 and 0.25, 0.9 and 0.1 and 2/3 and 1/3. The results are consistent and available upon request.

<sup>&</sup>lt;sup>13</sup>The results of probit models are not presented here and are available upon request. The estimates are consistent with those in logit model.

<sup>&</sup>lt;sup>14</sup>The number of resamples is set to 1000 for all models involved in this study.

<sup>&</sup>lt;sup>15</sup>The reason why the significance the estimates in each stock is not used is that  $Y_t$  is the discrete time series with auto-correlation. Therefore, the estimated standard deviation of  $\beta$  problematic and cannot be used for inference. However, the estimated  $\beta$  is still unbiased, so that results of nonparametric tests are valid.

For the constrained sample, we fit  $WNN_{f,t}$ ,  $WNP_{f,t}$ ,  $WNN_{m,t}$  and  $WNP_{m,t}$  into expression (11) with ordered logit and probit models. The dependent variable used here is still  $Y_t$  and t falls in the constrained sample. The mean of estimated  $\beta$  is reported in Table 2. Since here Wilcoxon test and Bootstrap confidence interval still yield slightly different results, we focus on the significant estimates confirmed by both methods. It can be observed that both negative and positive macro-economic news have significant marginal effects. Estimated marginal effect of  $WNN_{m,t}$  is at around 0.003, whereas that of  $WNP_{m,t}$  is around -0.004. In addition,  $WNN_{f,t}$  is estimated to have a significant positive marginal effect a around 0.003.

In conclusion, apart from occurrences of news, number and sentiment of news have further significant marginal effects on the likelihood of higher volatility state. More specifically, negative news can increase the likelihood of higher volatility states, while positive news can reduce that chance to a slightly larger degree. In addition, the effects of macroeconomic and firm-specific news are similar when sentiment and number of news are considered.

## 5 CONCLUSION

This paper investigates the impact of intraday firm-specific and macroeconomic news sentiment on volatility states of stock return for the constituent stocks of Dow Jones Composite Average spanning from January 1, 2000 to December 31, 2011. We firstly propose a MRS-FIGARCH model to identify the states of and proxy the stock return volatility. Then, discrete choice model is used to estimate and demonstrate how sentiment of public news releases can influence volatility states of asset return. The results show that our proposed MRS-FIGARCH model outperforms the existing time-varying-intercept FIGARCH models with respect to the modelling of structural change and evaluation of model performance. By employing discrete choice models on the volatility estimated from MRS-FIGARCH model, we demonstrate that news sentiment has a significant effect on likelihood of higher volatility state.

Moreover, nonparametric tests including Wilcoxon signed rank test and Bootstrap percentile confidence interval are employed to test the marginal effect of news. We firstly show that occurrence of news can significantly affect the likelihood of higher volatility states, with the effects of macroeconomic news being greater. It is further demonstrated that when sentiment and number of news are also considered, macroeconomic news has similar effects as firm-specific news. In addition, negative news generally increase the likelihood of higher volatility states, while positive news can reduce that chance to a slightly greater degree.

The empirical results can provide useful information for professional traders who read newswires like *Dow Jones*. They spend a considerable amount of money on these information sources and emphasize the importance of speed and accuracy of news. Newswire messages represent much of the overall information that traders receive on a real-time basis. The intraday impact of these newswire messages will help traders anticipate the potential effects on the volatility of the assets that they are monitoring. Furthermore, given that asset volatility is an important input in dynamic hedging strategies and derivatives pricing models, the impact of news sentiment can be incorporated to improve the accuracy of these strategies and models (Ho et al., 2013).

## REFERENCES

- Baillie, R. T., T. Bollerslev, and H. O. Mikkelsen (1996). Fractionally integrated generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* 74(1), 3 30.
- Baillie, R. T. and C. Morana (2009). Modelling long memory and structural breaks in conditional variances: An adaptive figarch approach. *Journal of Economic Dynamics and Control 33*(8), 1577–1592.
- Belkhouja, M. and M. Boutahary (2011). Modeling volatility with time-varying figarch models. *Economic Modelling* 28(3), 1106–1116.
- Diebold, F. X. and A. Inoue (2001). Long memory and regime switching. *Journal of Econometrics* 105(1), 131–159.
- Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica* 57(2), 357–384.
- Ho, K.-Y., Y. Shi, and Z. Zhang (2013). How does news sentiment impact asset volatility? evidence from long memory and regime-switching approaches. North American Journal of Economics and Finance.
- Wilfling, B. (2009). Volatility regime-switching in european exchange rates prior to monetary unification. *Journal of International Money and Finance* 28(2), 240–270.