INDIAN INSTITUTE OF MANAGEMENT CALCUTTA

WORKING PAPER SERIES

WPS No. 689/ November 2011

Impact of information arrival on volatility of intraday stock returns

by

Ashok Banerjee
Professor, IIM Calcutta, Joka, Diamond Harbour Road, Kolkata 700104, India

&

Subhadeep Paul, Soham Hazra, Rahul Dalmia
P G student, Indian Institute of Technology, Kharagpur. India

Impact of information arrival on volatility of intraday stock returns

Ashok Banerjee*, Subhadeep Paul†, Soham Hazra†, Rahul Dalmia†

November 22, 2011

Abstract

In this empirical study we have considered the impact of information flow on the volatility of a particular stock using high frequency return and news data on the Eurostoxx 50 market. In addition to using volume as a proxy for information flow, we have included company specific announcements, to the conditional variance of the Generalized Autoregressive Conditional Heteroscedastic model (GARCH). For this purpose we have constructed five measures of the impact of public information flow in the market transforming commonly available news scores through different techniques such as linear and exponential decreasing weight, impact function etc. We have analyzed the behaviour of volatility, estimated by squared returns for the next 4 hours after arrival of a non overlapping news, having a significant impact on the firm's stock return. A significant impact of the information flow accessed by the news score coefficient is observed for majority of in our analysis. Furthermore, the inclusion of the news scores variable improves the overall model in the sense that it increases the likelihood value of the model. However we do not observe any significant change in the volatility persistence due to inclusion of our news variable.

1 Introduction

While the market return of a stock is difficult to predict, there are well established models to predict return volatility. It has been observed in early sixties of the last century (Mandelbrot, 1963) that stock market volatility exhibits clustering, where periods of large returns are followed by periods of small returns. Later popular models of volatility clustering were developed by Engle (1982) and Bollerslev (1986). The autoregressive conditional heteroscedastic

*Professor, Indian Institute of Management, Calcutta. Corresponding author, Email : ashok@iimcal.ac.in

†Master degree student, Indian Institute of Technology, Kharagpur.

The authors acknowledge the contribution of Nayan Gupta and Abhishek Bishoyi, students of Indian Institute of Technology, Kharagpur for this paper. The authors also wish to acknowledge the generous data support provided by Ravenpack and Prof. Gautam Mitra of The Centre for the Analysis of Risk and Optimisation Modelling (CARISMA), Brunel University.

(ARCH) models (Engle, 1982) and generalized ARCH (GARCH) models (Bollerslev, 1986) have been extensively used in capturing volatility clustering in financial time series (Bollerslev et al. 1992). Using data on developed markets, several empirical studies (Akgiray, 1989; West et al, 1993) have confirmed the superiority of GARCH-type models in volatility predictions over models such as the naive historical average, moving average and exponentially weighted moving average (EWMA). GARCH models can replicate the fat tails observed in many high frequency financial asset return series, where large changes occur more often than a normal distribution would imply. Financial markets also demonstrate that volatility is higher in a falling market than it is in a rising market. This asymmetry or leverage effect was first documented by Black (1976) and Christie (1982). Two most popular GARCH formulations for describing this asymmetry are Threshold GARCH model (Glosten et al 1993) and Exponential GARCH model (Nelson, 1991). Empirical results also show that augmenting GARCH models with information like market volume or number of trades may lead to modest improvement in forecasting volatility (Lamoureux and Lastrapes 1990, Brooks, 1998 Jones et al, 1994). Encouraged by the fact that conditional volatility can be conveniently handled, researchers attempted to model volatility conditioned on prior information. The information set used included prior volatility, returns and volume.

The volume-volatility relationship is quite well researched. There are two competing theories that explain the volume-volatility relationship: (a) information theory and (b) dispersion of beliefs theory (Shalen, 1993). The first theory includes Clark's (1973) mixtures of distribution hypothesis (MDH) and the sequential arrival of information model. MDH stipulates that the volume-volatility relation originates from a joint dependence on a common event or variable. Such variable is often interpreted as the rate of information flows to the market. Thus, MDH states that both the asset price and trading volume change contemporaneously in response to new information. The sequential arrival of information model states that market information is not available to all participants simultaneously. The information is disseminated sequentially to traders and uninformed traders cannot properly infer the presence of informed trading. Investors receive the information in groups and trading happens after each group receives and interprets the information. Thus, the sequential arrival of information model is consistent with both contemporaneous and lagged relation between volume and volatility. The second theory, dispersion of beliefs, asserts that heterogeneous traders attach different importance to a set of information. A greater dispersion of beliefs creates excess price volatility and volume relative to their equilibrium levels. This theory helps one compare how informed and uninformed traders react to information. Clark (1973) and Epps and Epps (1976) found a positive relation between variance of price changes and aggregate trading volume for futures and stocks. In the literature, volume is conveniently used as a proxy for information. There are some studies which attempted to use trading volume of different categories of investors (e.g, liquidity traders, arbitragers, general public etc.) to establish the dispersion of beliefs theory. Trade volume of arbitragers was considered as an indicator of informed trading and similarly the volume of general public was treated as indicator for uninformed trading. The recent studies in volume-volatility used high frequency data. Price and volume data are available on an intraday basis and hence change in intraday volume data was used as a proxy for arrival of news/information.

News analytics, in general, measure the relevance, sentiment, novelty, and volume of news (Leinweber and Sisk, 2011). There is a growing body of literature that argues that media influence investor sentiment, hence asset prices, and asset volatility (e.g., Tetlock 2007, Barber and Odean, 2008 and Da, Engleberg, and Gao, 2009). Such research using news data, called news analytics, has largely been made possible with the availability of electronic news and news sentiment scores provided by select vendors. Using number of news releases by Reuter's News Service per unit of time as a measure of public information, Berry and Howe (1994) showed that there is a positive, moderate relationship between public information and trading volume. While there are several studies (e.g., Zhang and Skiena, 2010, Moniz et al, 2011) which used news or news sentiment scores to predict stock returns, a few studies have actually used news to predict volatility. Kalev(2004) considered arrival rate of firm specific news (frequency) as a proxy for information flow and empirically showed a positive and significant effect on conditional variance of stock returns.

The present study is a striking departure from the previous studies in explaining how new information impacts market volatility. Instead of using trading volume, this study uses high frequency firm-specific news as proxy for information flow and investigates the impact of news on volatility of stock returns. Information in the financial markets can be categorized as textual (bankruptcies, mergers and acquisitions, alliances etc.) and numerical (exchange rate, interest rate, profit margin etc.). Numerical information is often used by traders in their trading models to exercise the real-time trading activities or to predict future market behaviour. However, on many occasions, vital information is captured in the textual news. The real challenge, therefore, is to examine the impact of news stories on price behaviour.

The rest of the paper is organized as follows. The next section deals with methodology for volatility estimates and news impact function, section 3 describes sample and basic properties of return and data, section 4 shows analyses of results followed by conclusions in section 5.

2 Methodology

2.1 GARCH model augmented with volume ratio

Let r_t, $t = 1, 2 \ldots n$ denote log of price relatives at an intraday time-point t, where n is the number of return observations. Generally, the conditional mean μ_t of such return series $\{r_t\}$ can be modeled using a simple time series model such as a stationary ARMA(m, n) model, i.e.,

$$r_t = \mu_t + \epsilon_t$$

$$\mu_t = \phi_0 + \sum_{i=1}^{m} \phi_i r_{t-i} - \sum_{j=1}^{n} \theta_j a_{t-j} \tag{2.1}$$

Here the shock (or mean-corrected return) $\epsilon_t$represents the shock or unpredictable return, and m, n are non-negative integers. The conditional variance, then, can be modeled in a GARCH (p, q) process as:

$$h_t = \alpha_0 + \sum_{i=1}^{p} \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j h_{t-j} \tag{2.2}$$

$$\epsilon_t = z_t \sqrt{h_t}$$

where $\alpha_0 > 0$,$\alpha_i \geq 0, \beta_j \geq 0, \sum_{i=1}^{max(p,q)} (\alpha_i + \beta_i) < 1$with $\alpha_i = 0$ for $i > p$ and $\beta_j = 0$ for $j > q$. Z_t is a white-noise with mean zero and variances one and h_t is the conditional variance of ϵ_t.

To understand the effect of traded volume on volatility, we have fitted simple GARCH (1, 1) model initially for conditional volatility augmenting it with the ratio of successive time period (1 minute) traded volumes which we call "volume ratio" ($\frac{v_t}{v_{t-1}}$), and one time period lagged ratio of successive trading volumes which we call "lagged volume ratio" ($\frac{v_{t-1}}{v_{t-2}}$).

We will use these ratios as independent variables in the conditional volatility equation:

$$h_t = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \beta_1 h_{t-1} + \gamma \log \frac{v_t}{v_{t-1}} \tag{2.3}$$

Where, v_t denotes trading volume on day t. Earlier empirical studies show a decrease in persistence along with significantly improved model when trading volume is included as proxy of the intensity of information arrival in the conditional variance equation of the GARCH model (Lamoureux & Lastrapes, 1990). We observed similar phenomenon with trading volume but the model with lagged volume, although being significanctly improved, fails to decrease the persistence well.

2.2 Weakness of traded volume as a proxy for information : model with news scores

Since volume and volatility impact each other and both are influenced by the news arrival, volatility can not be regressed with contemporaneous volume assuming it as an exogenous variable due to simultaneity bias (Lamoureux & Lastrapes, 1990). Kalev P S (2004) documents more such issues with using contemporary traded volume such as liquidity trading, heterogeneity among traders, reservation values, revision of their dispersed beliefs, and strategic trading to exploit an informational advantage by the informed traders. Moreover since our final aim is to forecast future volatility through appropriate GARCH model we can not use contemporaneous volume as an explanatory variable in our equations for the obvious reasons that the model in that case will be incapable of forecasting. So instead of using contemporaneous volume we include lagged volume (the "lagged volume ratio" we have defined earlier) in our model as an exogenous variable. The GARCH model is similarly extended by adding the news impact scores (N_t) to the conditional variance equation. Our hypothesis is that news scores as a direct measure of information intensity improve the model in explaining the stylized facts of volatility in financial time series.

$$h_t = \alpha_0 + \alpha_1\epsilon_{t-1}^2 + \beta_1 h_{t-1} + \gamma \log\frac{v_t}{v_{t-1}} + \delta N_t \tag{2.4}$$

Our aim in this paper is to observe the individual significance of the coefficient of news scores measure δ which would give the significance of the effect information intensity on volatility both in presence and absence of the lagged volume term. This will give some direction to whether we should include this term in our GARCH conditional variance equation to explain volatility. We will also be testing the overall significance of the model, increase in power of explaining market volatility dynamics and forecasting. We compare the models with the likelihood ratio test (LRT) or equivalently AIC criterion. The use of higher order GARCH model is deliberately avoided as there are enough empirical evidence that a simple GARCH (1,1) adequately fits many financial time series (Sharma et al., 1996).

2.3 Continuous news scores and impact as a function of time

RavenPack[1] provided news impact scores NIP (News impact projection) which is the impact of the news expected on volatility over the next 120 minutes (which we will call "impact interval") based on their market response methodology. The score takes values between 0 and 100. The higher the score, the higher the confidence that a story will have an impact in terms of stock volatility. Much like the news sentiment scores available in the market, NIP is also a discrete measure, one value for one news. But all other variables in the model, return, volume and volatility are continuous in time measures. Since we are only considering the highly relevant news for a company there are limited number of news and hence limited number of NIP points in any time interval. So we need some methods of transforming the discrete NIPs into continuous in time measures to use the news impact variable effectively in the model. We construct 5 different continuous in time measures of our NIP keeping in mind our objective of quantifying the appropriate impact of all relevant news on the volatility of the stock at a particular time point .

2.3.1 Continuous NIP measures

In the first method the same impact score (NIP of the news) is assigned for 120 minutes (within the same day) after the arrival of the news. If more than one news have impact on a particular time point news impact score for that time point is simply the sum of the NIPs. This method neither differentiates impact of the news with passage of time, (the impact is assumed to be same at any time point in the 120 minute interval following it) nor does it impose any penalty on the impact of a news if another news comes within the impact interval. We call this new continuous series news scores 1 (N_1) and the model with this measure as model 1.

[1]RavenPack is a leading provider of news analytics and machine-readable content. The company specializes in linguistic analysis of high volume, real-time news from high-end newswire services. RavenPack News Scores includes metrics derived from textual news stories for the purpose of representing their qualitative nature in a quantitative manner.

$$h_t = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \beta_1 h_{t-1} + \gamma \log \frac{v_t}{v_{t-1}} + \delta N_{1t}$$

$$N_{1t} = \sum_j NIP_j$$

where j is the timepoint when the news arrived such that $t - 120 < j \leq t$.

However it is logical to assume the impact of a news will vary with the passage of time after its arrival. We take into consideration this variation of news impact on volatility in our second method where the news scores are calculated by assigning linearly decreasing weight to the NIP of a news for the impact interval previously defined after the arrival of the news. As earlier if more than one news have impact on a particular time point, we add the news scores (instead of adding the NIPs ofcourse) of all such news items for that time point. We call this series news scores 2 (N_2). The reported NIP value for a news is the average of the news scores calculated in this method for that news. The time point just after the arrival of the news has the highest news score assigned which is in fact greater than the reported NIP value. We call this method model 2 which is obtained by replacing N_{1t} in the earlier equation by N_{2t}.

$$N_{2t} = \sum_j m_{t-j} NIP_j, t - 120 < j \leq t.$$

Where as before j is the timepoint of arrival of the news and $t - j$ is the time passed since then. m_{t-j} is a linearly decreasing function of time passed since the arrival of the news, with maximum value being for $j = 0$ i.e m_t and minimum being for $j = 120$ i.e m_{t-120}.

In the third method we have attempted to down-weight the impact of the previous news as soon as another news arrives by a factor $0 < q < 1$. So as more and more news comes within the impact interval of a particular news, it gets exponentially decreasing weight. The news score at a particular time stamp is the sum of the news scores for all news having impact at that time period as before. A number of possible values for q is considered. We have chosen 0.7 as its value for our analysis on the basis of likelihood criterion. We call this construction news scores 3 (N_3) and the method, model 3.

$$N_{3t} = \sum_j q^{k_j} NIP_j, t - 120 < j \leq t$$

where k_j is the number of news that have arrived since the arrival of the news at timepoint j.

In our fourth construction we combine the ideas of the second and the third methods i.e we both weight the NIPs with a linearly decreasing weight function and penalize older news exponentially. We call this construction news scores 4 (N_4) and the method, model 4.

$$N_{4t} = \sum_j m_{t-j} q^{k_j} NIP_j, t - 120 < j \leq t$$

In our final construction method we attempt to estimate the weight function empirically rather than arbitrarily assigning a linear function. The construction method of this weight function which we call "news impact function" is outlined in the following subsection. We call this construction news scores 5 (N_5) and the method, model 5.

$$N_{5t} = \sum_j f(t-j) NIP_j, t-120 < j \leq t.$$

where $f(t-j)$ is the news impact function.

2.3.2 News impact function and graphical representations

To estimate a function of impact of a news on volatility over a 2 hour period empirically we look at the behavior of volatility using squared returns as a proxy over the next 180 minutes after the arrival of a particular news. For this purpose we have considered 30 distinct news with considerable high NIP values for each stock under study such that there is no news in 120 minutes intervals after and before the news under consideration, which would justify our argument that volatility movement in the time horizon considered is due to the news under consideration. When we plot the squared returns with time the "news impact curves" obtained has several common features across all stocks. To understand nature of this dependence we smooth the curves with robust locally weighted regression(Cleveland 1988) with smoothing parameter set at 0.50. Heavy smoothing brings out the common pattern in the curves- all of them have two smooth peaks, one immediately after the arrival of the news and the other some what lower peak after around 2 hour, more specifically (122.84±32.57) minutes. In between the peaks volatility decays continuously quadratically to a crest at around one hour or (76.47±24.58) minutes before moving up again to the second peak. The mean of the difference between these two peaks is observed to be around 45 minutes.The remarkable similarity between the graphs can easily be seen from the table(Appendix 2, table 13) and actual plots presented in Appendix 2.

The function obtained from this method which we call "impact function" is used to weight the discrete NIP measures to create a continuous NIP measure. Here it is important to differentiate between the impact function we are using and the "impact curve" of Engle and Ng (1993). In the later case its a curve which relates past return shocks (news) with current volatility which is basically a measure of how new information is incorporated theoretically into the volatility estimates of ARCH series of models. The equation of the impact curve in Engle and Ng is $h_t = A + \alpha\varepsilon_{t-1}^2$ where $h_t$is the conditional variance at time t and $\varepsilon_{t-1}$is the unpredictable return at time $t-1$. Clearly it brings out the assymetric effects of shocks (news) on volatility. The impact function we have constructed in this article is essentially an empirical function which measures the variation of intra day volatility (squared returns) minutes after the arrival of a news and is computed for the purpose of news scores. The observed behaviour of the squared returns can be explained with the Sequential arrival of information hypothesis, which states that information is reached to different traders and market players at different time lags and they are interpreted in markedly different ways.Informed strategic traders with heterogeneous information will react to the same news at different times to maximize his profit. As the news reaches the market, informed traders make the first few moves and the stocks volatility remains higher due to more activity which gradually settles down, which correspond to our first down slope in the figure 1 and settles down to a lower volatility level. A second series of traders,who have received the information much later then utilizes the information and trades resulting in the second surge in volatility of the stock which corresponds to our first up slope and reaches the second but generally lower peak before gradually decreasing again and going back to the normal level.

3 Sample

We have collected intraday price, volume and news information for the 50 companies that featured in Eurostoxx 50 during 2 February 2005 to 25 June 2008. As the focus of the study was on impact of news on volatility of stock returns, the primary filtering criterion used was the availability of news. Companies which did not have sufficient number of news items/scores during the above-mentioned period, were removed from our final sample. In our research , we have finally used intraday minute-wise return, volume, and news data for 20 companies of Eurostoxx 50. Our news scores were generated from a News Impact Scores, NIP provided by Ravenpack by methods which are documented later in the paper. If news is not available for the entire time horizon under study for a particular company we have considered the time span for which news data are available as our time horizon provided its not too short & there are ample news data available in that span. Only the news which are highly 'relevant' for the company are selected for analysis,If a news is reported multiple times only the first report of the news is considered where as others are discarded from analysis.

3.1 Return data

Intraday stock return time series introduces the added difficulty of being a discontinuous series. Although intra day log-returns are by definition calculated with closing prices of successive minutes, however, the opening minute of a day (09:00 CET-09:01 CET) has no actual 'previous minute' closing price. So, we have proxied the 'opening price of the day' as the closing price of the 'previous' minute for the opening minute. The time between previous day's closing to this day's opening is considered a single 'minute' (which has a closing price identical as opening price of this day). This takes care of the continuity of the time-series and minimizes any potential bias. A similar problem is obtained while we merge the return data with news data, while there is no trading on public holidays and weekends, information continues to flow. Since the return series is continuous, it could affect the relationship between return and information arrival. In our analysis we have considered only those news that have arrived in the trading hours and since opening minute return is calculated from the closing and opening prices of the opening minute only, a potential error could be avoided. For volume-ratio series, we have followed the similar method as price return.

The descriptive statistics given in the table 1 in Appendix1 shows the stylized facts of the high frequency return series. High kurtosis is observed in almost all the stocks considered indicating significant departure from normality of the series. A table of volume traded in each quarter for all the stocks are also provided(table 2, Appendix 1) which helps us to identify the most traded stocks as well as a possible relationship between volume traded and information arrival of stocks.

3.2 News data

The only news variable in this study, is the NIP score (news impact) as reported by Ravenpack. According to RavenPack manuals and white papers, it determines a classification rule from a large sample of news headlines on the basis of the words and word combinations of the news and its impact on the market. With this rule different advanced machine learning methods are applied to the news database for creating the impact scores that identifies the probability that the volatility of a particular stock to be higher or lower than the volatility of the market.The technique is tested and refined on a large number of different news stories.

For our analysis,only the news which are highly relevant for the company have been selected. If a news is reported multiple times only the first report of the news is considered where as others are discarded from analysis. Since our main aim is to analyze the impact of news on stock volatility, the return and the news data were merged together into a single database. Hence we have only considered those stocks for which sufficient simultaneous news and return data were available. The number of news in each quarter for all the stocks considered has been presented in table 3, Appendix 1.

4 Empirical Results

Our empirical results reveal there is a significant impact of public information flows on conditional volatility of majority of stock returns. This impact is better evident for those companies which have large amount of news data within the interval considered. The estimation of GARCH(1,1) model with continuous news scores as a news variable in the conditional variance equation shows a statistically significant coefficient (delta) at 0.01 level with at least one news score measure for most of the stocks under consideration. The log likelihood value also increases significantly after the inclusion of the news scores in the GARCH equation and the log likelihood ratio test (LRT) rejects the NULL hypothesis of no improvement in model at 0.05 and 0.01 significance levels which means that the model after inclusion of the NIP term does better in explaining the volatility variations. Table 5 in Appendix gives the log likelihood values and the LRT test statistic along with its p-value for different models. All changes of log likelihood over the GARCH(1,1) model are statistically significant at <0.0001 level as evident from Table 6. The LR test clearly reveals the superiority of the model with news scores,which indicates that volatility of the stocks can be better explained with the knowledge of the news variable. Table 7 gives the coefficients of simple GARCH(1,1). Table 8 shows the coefficients with augmenting lagged-volume. The volume coefficients are found to be significant. In Table 9 through Table 13, the coefficients of the GARCH model 1 to 5 augmented with different news score measures as mentioned before, both with and without presence of volume are presented. The estimated lagged volume coefficient is always statistically significant both in presence or absence of the news scores term. The corresponding coefficients for the news scores are found to be statistically significant for most of the companies, even in the presence of the volume coefficient, which indicates the impact of public information on volatility of the stocks.

5 Conclusions

In this paper we have investigated the impact of public information on volatility of stocks using high frequency return data and news scores. Our analysis reveals two important empirical results. The first one is that the information variable proxied as news scores in this paper significantly affects the volatility of the stock. This is evident from the statistically significant estimates of the news score coefficient in majority of the companies for at least one news score measure as well as significant value of the LRT statistic. However the volatility persistence does not appear to decrease significantly with the addition of news scores. This is expected since we have considered discrete news arrival events and the possible impact of that news on the volatility of the stock instead of taking amount of news arrival, there is no correlation structure present among our news scores. The second result is the characterization of the impact of a news on volatility over time immediately after the arrival of the news through the impact function obtained from an analysis of squared returns post news arrival. A distinct common structure for the function has been obtained which is independent of the stock. Further research can reveal more interesting features of this function. Despite limitations of small amount of news data compared to return data for most of the companies, results obtained in our study emphasizes the need of revising the GARCH model with information flow variable for building a more efficient model for volatility forecasting.

6 Appendix 1

In the following tables blank cells indicate the coefficient value was not available as the numerical process did not converge to a solution. Results which are significant at the 5 % statistical significance level have been highlighted.

Table 1: Summary statistics of return

company	Count	Mean($\times 10^{-7}$)	Std dev	Skewness	Kurtosis	Max	Min
Deutsche bank	421942	-5.21	0.000685	-0.0095	16.3222	0.0213	-0.0207
Sap	413888	6.42	0.000735	-0.0456	27.9019	0.0226	-0.0301
Deutsche Telecom	430236	-6.75	0.000709	-0.1602	15.0432	0.0188	-0.0238
Volkswagen	396509	55.4	0.000758	0.1912	22.3825	0.0269	-0.0215
Eni	422927	-9.14	0.000575	-0.0448	17.9148	0.0203	-0.0159
Iberdrola	376860	-2.39	0.001038	0.0767	11.4126	0.0287	-0.0185
Telefonica	435407	11.2	0.000686	0.0957	12.3043	0.0227	-0.0190
Alcatel-Lucent	188802	-46.7	0.001450	-0.1698	13.2874	0.0259	-0.0407
France Telecom	165980	-16.3	0.000866	-0.0703	22.1502	0.0204	-0.0266
Enel	421080	-3.31	0.000621	-0.0261	31.9952	0.0299	-0.0299
Air liquide	125073	21.4	0.000938	0.0040	12.4606	0.0195	-0.0207
Bayer	406874	11.8	0.000798	-0.0005	22.1930	0.0259	-0.0217
Bbva	423627	7.62	0.000701	-0.0370	8.8737	0.0162	-0.0149
Daimler	110382	-33.9	0.000984	0.1362	20.2752	0.0291	-0.0209
Deutsche borse	350096	8.84	0.001016	0.0291	18.0148	0.0296	-0.0201
Generali asset	412951	-3.00	0.000620	0.6483	27.2989	0.0936	-0.0951
Intesa sanpaolo	181759	-17.4	0.000868	0.0755	33.9084	0.0280	-0.0262
Muench rueckvers	402692	-6.39	0.000683	-0.0422	27.3562	0.0239	-0.0291
Nokia	338655	19.9	0.000784	-1.4904	418.3064	0.0471	-0.0726
Repsol	407627	3.74	0.000868	-0.0592	46.0818	0.0346	-0.0291
Rwe	407649	13.4	0.000736	-0.0629	13.4522	0.0208	-0.0198
Santander	429542	5.59	0.000816	0.0899	10.7170	0.0269	-0.0221
Siemens	427891	-3.92	0.000715	0.1073	33.2217	0.0365	-0.0188
Suez	95091	26.6	0.001004	0.3178	30.6202	0.0317	-0.0237
Unicredit	51522	-34.1	0.000971	-0.3395	8.9540	0.0114	-0.0171
Unilever	161500	-1.27	0.000816	0.4309	52.2967	0.0384	-0.0301
Saint-Gobain	390567	-24.7	0.000907	0.0709	21.7997	0.0293	-0.0323
Sanofi-Aventis	419726	-6.67	0.000803	0.3261	23.8643	0.0354	-0.0189
Societe Generale	408931	-11.6	0.000911	0.1106	24.8932	0.0300	-0.0235
Total	432852	-2.85	0.000735	0.3096	27.8660	0.0408	-0.0152
Vivendi	418936	-13.1	0.000800	0.5316	56.9133	0.0479	-0.0264

Table 2: volume traded in each quarter ($\times 10^8$)

company	quarter no														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	Total
deutsche bank	1.55	2.46	2.52	2.28	2.64	3.23	2.44	2.44	3.05	3.01	5.28	3.84	5.35	3.90	43.99
sap	2.35	4.44	3.16	3.08	4.18	4.37	5.22	4.32	6.48	6.18	6.10	5.23	6.06	4.04	65.22
deutsche telecom	10.19	15.36	13.85	14.56	18.55	22.90	18.58	16.59	20.47	20.70	17.54	18.71	25.49	16.90	250.41
volkswagen	1.01	1.43	2.64	1.69	1.90	1.46	1.37	2.78	4.51	1.26	1.50	1.33	1.58	0.88	25.32
eni	9.57	24.40	17.63	25.15	13.11	22.94	13.05	20.85	16.53	23.40	18.12	23.51	14.44	19.77	262.46
iberdrola	6.97	11.09	11.07	9.89	12.90	10.95	11.75	14.40	18.29	24.85	16.83	16.40	26.80	18.16	210.36
telefonica	14.66	19.94	23.20	29.97	27.62	28.28	21.94	28.53	29.55	27.94	30.08	25.49	25.82	21.18	354.21
alcatel-lucent								4.46	13.11	10.56	15.29	11.91	14.46	10.56	80.34
france telecom									2.70	8.21	9.07	7.85	9.56	8.64	46.03
enel	14.37	28.67	26.64	38.49	28.27	32.07	17.69	35.97	30.49	33.81	26.54	33.64	27.78	35.98	410.41
air liquide										0.19	0.89	0.57	0.77	0.62	3.03
bayer	1.95	2.54	2.47	2.29	3.53	4.44	3.26	2.83	3.47	3.71	3.72	3.12	4.49	2.53	44.35
bbva	8.10	12.04	12.56	11.74	13.37	13.22	10.99	13.89	19.68	19.16	21.36	20.00	23.21	17.21	216.52
daimler											3.94	4.99	7.10	5.14	21.16
deutsche borse	1.41	1.49	0.96	0.78	1.19	1.50	0.77	1.20	1.26	1.04	1.29	0.99	1.54	1.44	16.86
generali ass	2.74	6.17	4.58	4.81	5.91	6.42	3.76	5.66	5.58	6.13	6.30	6.33	5.72	4.46	74.57
intesa sanpaolo									54.48	76.99	52.30	49.68	46.52	62.27	342.23
muench rueckvers	0.58	1.04	1.29	1.41	1.32	1.24	1.18	1.16	1.76	1.48	1.38	1.15	1.77	1.00	17.75
nokia	7.61	14.89	17.01	14.03	15.20	15.19	13.78	13.59	18.25	14.64	14.67	12.82	6.93	5.44	184.06
repsol	3.19	5.00	4.97	4.44	6.03	5.40	4.55	8.01	5.40	5.55	5.32	5.69	6.84	5.29	75.69
rwe	1.57	1.72	1.96	1.97	1.81	2.13	1.73	2.19	2.49	2.49	2.45	2.11	3.14	2.15	29.92
santander	15.79	22.70	24.64	23.33	23.45	24.78	19.43	27.23	34.37	35.80	31.10	38.29	43.56	32.07	396.54
siemens	2.03	3.01	3.11	3.07	3.24	4.14	2.95	3.34	4.26	5.24	5.16	3.91	5.42	3.53	52.41
suez											0.43	3.35	3.98	2.77	10.52
unicredit													70.30	111.65	181.95
unilever									3.59	5.92	6.62	6.15	7.41	5.57	35.25

* 'blank' indicates no data available for that quarter.

Table 3: Number of relevent news in each quarter

company	quarter no														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	Total
deutsche bank	339	369	335	398	460	415	326	378	431	410	439	312	409	414	5435
sap	115	135	119	111	136	135	165	119	193	123	153	143	134	119	1900
deutsche telecom	164	179	259	224	234	277	175	298	188	196	129	127	148	164	2762
volkswagen	143	168	397	308	333	253	304	472	401	194	223	282	251	280	4009
eni	73	101	69	102	248	74	142	174	116	208	133	189	123	168	1920
iberdrola	29	45	46	25	65	27	53	135	73	83	36	40	169	80	906
telefonica	121	171	141	262	245	146	120	89	216	112	134	136	71	94	2058
alcatel-lucent	-	-	-	-	-	-	-	46	214	172	187	167	173	137	1096
france telecom	-	-	-	-	-	-	-	-	26	125	93	113	83	199	639
enel	141	160	99	67	253	103	66	95	309	211	136	140	115	132	2027
air liquide	-	-	-	-	-	-	-	-	-	5	54	37	28	19	143
bayer	62	103	122	208	222	255	167	172	131	121	118	128	115	79	2003
bbva	106	152	93	80	91	76	77	141	120	129	73	59	48	45	1290
daimler	-	-	-	-	-	-	-	-	-	-	126	209	208	201	744
deutsche borse	129	157	92	80	113	299	109	227	161	155	86	80	105	117	1910
generali ass	42	67	63	46	97	84	61	57	68	54	54	88	72	44	897
intesa sanpaolo	-	-	-	-	-	-	-	-	67	143	74	63	91	67	505
muench rueckvers	88	103	144	122	102	67	45	82	98	104	78	160	88	59	1340
nokia	254	277	353	347	435	397	316	335	383	309	337	430	133	99	4405
repsol	44	83	68	55	153	85	37	71	55	43	49	61	57	38	899
rwe	75	69	90	129	132	145	166	224	176	164	156	178	229	152	2085
santander	63	28	114	118	89	104	105	95	112	85	76	108	89	110	1296
siemens	113	212	216	176	156	209	136	168	232	303	301	287	247	216	2972
suez	-	-	-	-	-	-	-	-	-	-	5	55	36	44	140
unicredit	-	-	-	-	-	-	-	-	-	-	-	-	156	153	309
unilever	-	-	-	-	-	-	-	-	2	3	9	12	6	3	35

* - indicates no data available for that quarter.

Table 4: Change in the log likelihood(D-stat) after adding lagged volume & different measures of news scores in the conditional variance of GARCH(1,1) model compared to the simple GARCH(1,1) model. (The first column gives the log likelihood of simple GARCH and the following 6 columns give the changes in that for different models.)

Company	log likelihood of simple GARCH	lagged volume	model 1	model 2	model 3	model 4	model 5
Alcatel-Lucent	805334	994	1024	1034	1034	1034	1024
Bbva	903725	664	664	672	664	668	672
Deutsche Telecom	912084	422	432	434	434	434	432
Deutsche Borse	869807	936	1256	1278	1008	998	1258
Enel	913627	920	928	930	926	928	928
Eni	935783	708	710	708	710	710	708
france telecom	864164	662	664	664	662	664	664
Generali asset	942404	572	572	572	572	572	572
iberdrola	841783	886	886	888	886	892	886
Intesa Sanpaolo	861068						
Nokia	888027	300	410	430	582	422	560
RWE	900745	106	134	128	112	110	116
SAP	916920	220	230	230	224	230	232
Telefonica	907825	566	566	566	568	568	566
vow	888991	280	302	294	290	298	306
Bayer	890209	187	187	187	188	188	187
Muench rueckvers	914142	172	172	172	172	172	172
Repsol	901400	297	297	297	297	297	297
Santander	877095	340	337	337	343	344	343
Siemens	928513	134	148	146	164	159	149

Table 5: Change in the log likelihood(D-stat) along with the statistical significance (p value) after adding different measures of news scores to the GARCH(1,1) augmented with lagged volume model.

Company	news score1	news score 2	news score 3	news score 4	news score 5
Alcatel	**30 (<0.001)**	**40 (<0.001)**	**40 (<0.001)**	**40 (<0.001)**	**30 (<0.001)**
Bbva	0	**8 (0.0047)**	0	**4 (0.0455)**	**10 (0.0016)**
Deutsche Telecom	**10 (0.0016)**	**12 (<0.001)**	**12 (<0.001)**	**12 (<0.001)**	**10 (0.0016)**
Deutsche Borse	**320 (<0.001)**	**342 (<0.001)**	**72 (<0.001)**	**62 (<0.001)**	**322 (<0.001)**
Enel	**8 (0.0047)**	**10 (0.0016)**	**6 (0.0143)**	**8 (0.0047)**	**8 (0.0047)**
Eni	2 (0.1573)	0	2 (0.1573)	2 (0.1573)	0
france telecom	2	2	0	2	2
Generali asset	0	0	0	0	0
iberdrola	0	2 (0.1573)	0	**6 (0.0143)**	0
Intesa sanpaolo	**8(0.0047)**	**8(0.0047)**	**4(0.0455)**	**4(0.0455)**	**8(0.0047)**
Nokia	**110 (<0.001)**	**130 (<0.001)**	**282 (<0.001)**	**122 (<0.001)**	**230 (<0.001)**
RWE	**28 (<0.001)**	**22 (<0.001)**	**6 (0.0143)**	**4 (0.0455)**	**10 (0.0016)**
SAP	**10 (0.0016)**	**10 (0.0016)**	**4 (0.0455)**	**10 (0.0016)**	**12 (<0.001)**
Telefonica	0	0	2 (0.1573)	2 (0.1573)	0
vow	**22 (<0.001)**	**14 (<0.001)**	**10 (0.0016)**	**18 (<0.001)**	**26 (<0.001)**
Bayer	0	0	2 (0.1573)	2 (0.1573)	0
Muench rueckvers	0	0	0	0	0
Repsol	0	0	0	0	0
Santander	**6 (0.0143)**	**6 (0.0143)**	**6 (0.0143)**	**8 (0.0047)**	**6 (0.0143)**
Siemens	**28 (<0.001)**	**24 (<0.001)**	**60 (<0.001)**	**50 (<0.001)**	**30 (<0.001)**

Table 6: Coefficients for GARCH(1,1)

company	alpha	beta	persistence
Alcatel	0.0602	0.9330	0.9932
Bayer	0.1088	0.8621	0.9709
BBVA	0.0743	0.9091	0.9834
Deutsche Telecom	0.0652	0.8957	0.9609
Deutsche Borse	0.1827	0.8003	0.9830
ENEL	0.0449	0.9314	0.9762
ENI	0.0507	0.9342	0.9849
France_Telecom	0.0776	0.9073	0.9849
Generali asset	0.0740	0.9052	0.9792
Iberdrola	0.1139	0.8347	0.9486
Intesa_Sanpaolo	0.0414	0.9479	0.9893
Muench_Rueckvers	0.1174	0.8706	0.9880
Nokia	0.0952	0.9030	0.9983
Repsol	0.1107	0.8637	0.9744
RWE	0.1054	0.8750	0.9803
Santander	0.0773	0.9098	0.9871
SAP	0.1014	0.8750	0.9764
Siemens	0.1060	0.8790	0.9850
Telefonica	0.0725	0.9041	0.9767
Vow	0.1258	0.8551	0.9809

Table 7: Coefficients for GARCH(1,1) along with lagged volume

company	alpha	beta	persistence	volume coefficient($\times 10^{-8}$)
Alcatel	0.052712	0.940578	0.99329	4.81 (<0.0001)
Bayer	0.099593	0.872675	0.972268	1.812 (<0.0001)
BBVA	0.07418	0.906984	0.981164	1.089 (<0.0001)
Deutsche Telecom	0.058555	0.908429	0.966984	1.205 (<0.0001)
Deutsche Borse	0.169144	0.81451	0.983654	2.074 (<0.0001)
ENEL	0.0366	0.943981	0.980581	1.68 (<0.0001)
ENI	0.0456	0.940222	0.985822	9.831 (<0.0001)
France_Telecom	0.070983	0.914246	0.985229	2.246 (<0.0001)
Generali asset	0.067758	0.91192	0.979678	0.8058 (<0.0001)
Iberdrola	0.110549	0.843402	0.953951	2.887 (<0.0001)
Intesa_Sanpaolo	0.083752	0.785739	0.869491	3.941 (<0.0001)
Muench_Rueckvers	0.112541	0.87467	0.987211	0.9912(<0.0001)
Nokia	0.091476	0.906397	0.997873	1.052 (<0.0001)
Repsol	0.105093	0.869951	0.975044	1.279 (<0.0001)
RWE	0.099446	0.881765	0.981211	1.344 (<0.0001)
Santander	0.076084	0.908219	0.984303	1.183 (<0.0001)
SAP	0.094031	0.883855	0.977886	1.102 (<0.0001)
Siemens	0.098943	0.885651	0.984594	0.8208 (<0.0001)
Telefonica	0.070497	0.906316	0.976813	1.125 (<0.0001)
VOW	0.119415	0.860668	0.980083	1.305 (<0.0001)

Table 8: Coefficients for news score 1 in model 1, both with and without lagged volume

company	alpha	beta	volume coefficient($\times 10^{-8}$)	News score in model 1 with volume($\times 10^{-11}$)	News score in model 1 without volume($\times 10^{-11}$)
Alcatel	0.052911	0.940173	4.766 (<.0001)	**2.28 (<.0001)**	**3.36(<0.001)**
Bayer	0.099925	0.872368	1.798 (<.0001)	-3.66 (0.7791)	-3.58(0.8030)
BBVA	0.074026	0.906477	1.125 (<.0001)	-13.3 (0.4234)	8.16(0.6428)
Deutsche Telecom	0.058988	0.907003	1.192 (<.0001)	**0.441 (0.0035)**	**0.498(0.004)**
Deutsche Borse	0.164843	0.81812	2.051 (<.0001)	**603.9 (<.0001)**	**633.5(<0.001)**
ENEL	0.036955	0.943185	1.675 (<.0001)	**0.3(0.0020)**	**0.436(0.0004)**
ENI	0.045432	0.94045	9.853 (<.0001)	0.062 (0.3741)	0.084(0.2795)
France Telecom	0.070654	0.914605	2.245 (<.0001)	-0.108 (0.7472)	0.13(0.7327)
Generali asset	0.067725	0.91198	8.067 (<.0001)	-6.64 (0.2057)	-3.4(0.5834)
Iberdrola	0.110927	0.84303	2.907 (<.0001)	-0.016 (0.9861)	-0.502(0.6250)
Intesa sanpaolo					**60.6(0.0105)**
Muench Rueckvers	0.112351	0.874837	0.9891 (<.0001)	-3.59 (0.4589)	-5.87(0.2872)
Nokia	0.00278	0.244452	0.524(<0.001)	**13.20 (<0.0001)**	**37.2(<0.0001)**
Repsol	0.104684	0.870774	1.281 (<.0001)	8.1 (0.3255)	5.84(0.444)
RWE	0.099447	0.881327	1.344 (<.0001)	**138.3(<.0001)**	**150.5(<0.0001)**
Santander	0.080207	0.903671	1.17 (<0.001)	**-2.93(<0.001)**	-2.27(0.3619)
SAP	0.094156	0.883594	1.105 (<.0001)	**0.614 (0.0012)**	**0.673(0.0009)**
Siemens	0.098868	0.885333	0.8224 (<.0001)	**-0.696 (<.0001)**	**-0.807(<0.0001)**
Telefonica	0.075847	0.897832	1.13 (<0.001)	**0.232(0.001)**	**0.318(0.0321)**
Vow	0.119696	0.859974	1.302E-8 (<.0001)	**0.815 (<.0001)**	**0.925(<0.0001)**

Table 9: Coefficients for news score 2 in model 2, both with and without lagged volume

company	alpha	beta	volume coefficient($\times 10^{-8}$)	News score in model 2 with volume($\times 10^{-11}$)	News score in model 2 without volume($\times 10^{-11}$)
Alcatel	0.052911	0.940161	4.795 ($<$.0001)	**2.56 ($<$.0001)**	**3.65($<$0.0001)**
Bayer	0.099797	0.872424	1.805 ($<$.0001)	1.48 (0.9113)	-2.47(0.8650)
BBVA	0.071369	0.910923	1.07 ($<$.0001)	-1.97 (0.9061)	9.74(0.5782)
Deutsche Telecom	0.058834	0.907255	1.20 ($<$.0001)	**0.481(.0035)**	0.503(0.0033)
Deutsche Borse	0.164286	0.81872	2.05 ($<$.0001)	**601.2($<$.0001)**	641.5 ($<$0.0001)
ENEL	0.036635	0.943955	1.674 ($<$.0001)	**0.292 (0.0018)**	**0.436 (0.0002)**
ENI	0.045751	0.940087	0.9827 ($<$.0001)	0.0559 (0.4237)	0.064 (0.4071)
France_Telecom	0.070754	0.91449	2.251 ($<$.0001)	-0.128 (0.6972)	0.101 (0.7871)
Generali Asset	0.067702	0.912024	0.807 ($<$.0001)	-5.88 (0.3525)	-5.4 (0.4057)
Iberdrola	0.111329	0.841951	2.872 ($<$.0001)	-1.36 (0.1300)	-1.19 (0.2238)
Intesa Sanpaolo					**51.8 (0.0189)**
Muench_Rueckvers	0.112003	0.875041	0.9923 ($<$.0001)	-9.1 (0.8869)	-7.15 (0.2597)
Nokia	0.002842	0.329624	0.561($<$.0001)	**13.9 ($<$.0001)**	**37.2 ($<$0.0001)**
Repsol	0.104868	0.870281	1.28 ($<$.0001)	7.43 (0.4316)	5.76 (0.5254)
RWE	0.099162	0.881793	1.352 ($<$.0001)	**118.9 (0.0002)**	**130.3 ($<$0.0001)**
Santander	0.080388	0.903451	1.161 ($<$.0001)	-3.08 (0.2213)	-2.3 (0.3622)
SAP	0.09418	0.883593	1.102 ($<$.0001)	**0.578 (0.0016)**	**0.617(0.0018)**
Siemens	0.098819	0.88544	0.8223 ($<$.0001)	**-0.628 ($<$.0001)**	**- 0.731($<$0.0001)**
Telefonica	0.071437	0.899524	1.23	0.346	0.232 (0.1117)
Vow	0.119958	0.859485	1.299 ($<$.0001)	**1.04($<$.0001)**	**1.18 ($<$0.0001)**

Table 10: Coefficients for news score 3 in model 3, both with and without lagged volume

company	alpha	beta	volume coefficient($\times 10^{-8}$)	News score in model 3 with volume($\times 10^{-11}$)	News score in model 3 without volume($\times 10^{-11}$)
Alcatel	0.098864	0.878359	4.504 ($<$.0001)	-1.76 (0.1495)	5.31($<$0.0001)
Bayer	0.09986	0.872494	1.799 ($<$.0001)	-38.2 (0.3140)	-58.0(0.1645)
BBVA	0.074299	0.906555	1.099 ($<$.0001)	5.28 (0.9064)	39.6(0.3243)
Deutsche Telecom	0.058946	0.907307	1.19 ($<$.0001)	**0.944 (0.0007)**	**1.00 (0.0017)**
Deutsche Borse	0.166791	0.816756	2.06 ($<$.0001)	**1052 ($<$.0001)**	**1184($<$0.0001)**
ENEL	0.036905	0.943236	1.665 ($<$.0001)	**0.474 (0.0067)**	**0.478(0.0007)**
ENI	0.0455	0.9403	0.985($<$0.001)	-0.167(0.1565)	-0.17(0.196)
France_Telecom	0.07085	0.914335	2.247 ($<$.0001)	-0.0743 (0.8855)	0.57 (0.3503)
Generalli_Ass	0.067805	0.911775	0.8041 ($<$.0001)	-12.7 (0.3480)	-8.94 (0.54)
Iberdrola	0.110701	0.843049	2.871 ($<$.0001)	-1.08 (0.4206)	-2.58 (0.0902)
Intesa Sanpaolo					**103.9 (0.0984)**
Muench_Rueckvers	0.111978	0.875116	1.006 ($<$.0001)	38.4 (0.4520)	-4.42(0.929)
Nokia	0.082872	0.914466	1.016 ($<$.0001)	**69.7 ($<$.0001)**	**74.9 ($<$0.0001)**
Repsol	0.104799	0.870365	1.281 ($<$.0001)	33.9 (0.6014)	2.62 (0.97)
RWE	0.09909	0.881959	1.348 ($<$.0001)	**197.7 (0.0515)**	**191.4 (0.0771)**
Santander	0.080186	0.903918	1.133 ($<$.0001)	**-16.9 (0.0012)**	**-17.5 (0.0007)**
SAP	0.094055	0.883863	1.104 ($<$.0001)	**0.691 (0.0203)**	**0.688 (0.0354)**
Siemens	0.099111	0.884548	0.8199 ($<$.0001)	**-1.85 ($<$.0001)**	**-2.15 ($<$0.0001)**
Telefonica	0.074497	0.898476	1.16 ($<$.0001)	**1.34(0.02)**	**0.531 (0.0432)**
Vow	0.11978	0.860102	1.311 ($<$.0001)	**1.13 (0.0027)**	**1.08 (0.0056)**

Table 11: Coefficients for news score 4 in model 4, both with and without lagged volume

company	alpha	beta	volume coefficient($\times 10^{-8}$)	News score in model 4 with volume($\times 10^{-11}$)	News score in model 4 without volume($\times 10^{-11}$)
Alcatel	0.052937	0.940101	4.762 (<.0001)	**3.88 (<.0001)**	**5.44(<0.0001)**
Bayer	0.099941	0.872314	1.804 (<.0001)	-32.4 (0.3585)	-50.5(0.1902)
BBVA	0.073109	0.908282	1.088 (<.0001)	17.1 (0.6434)	46.4(0.2217)
Deutsche Telecom	0.058826	0.907398	1.194 (<.0001)	**0.824 (0.0009)**	**0.879 (0.0023)**
Deutsche Borse	0.166495	0.816978	2.065 (<.0001)	**869.6 (<.0001)**	**981 (<0.0001)**
ENEL	0.03729	0.942572	1.673 (<.0001)	**0.458 (0.0054)**	**0.7(0.0006)**
ENI	0.045507	0.940343	0.9854 (<.0001)	-0.114 (0.2992)	-0.135 (0.2706)
France_Telecom	0.071186	0.913897	2.253 (<.0001)	-0.205 (0.6722)	0.38 (0.4847)
Generali_Asset	0.067703	0.912013	0.8064 (<.0001)	-8.52 (0.6074)	-1.92 (0.918)
Iberdrola	0.111034	0.842451	2.855 (<.0001)	**-3.3 (0.0102)**	**-3.39 (0.0155)**
Intesa Sanpaolo					**97 (0.0984)**
Muench_Rueckvers	0.112588	0.874624	0.9912 (<.0001)	21.6 (0.5912)	11.2 (0.785)
Nokia	0.087008	0.910761	1.033 (<.0001)	**32.6 (<.0001)**	**35.7 (<0.0001)**
Repsol	0.104836	0.870334	1.279 (<.0001)	44.2 (0.4546)	22.4 (0.7265)
RWE	0.099189	0.881821	1.344 (<.0001)	136.3 (0.1088)	133.5 (0.1472)
Santander	0.078319	0.905836	1.136 (<.0001)	**-12.1 (0.0057)**	**-12.9 (0.0047)**
SAP	0.094055	0.883863	1.104 (<.0001)	**0.691 (0.0203)**	
Siemens	0.098933	0.88487	0.8204 (<.0001)	**-1.54 (<.0001)**	**-1.78 (<0.0001)**
Telefonica	0.069623	0.907198	1.13	-0.669	0.268 (0.2504)
Vow	0.119765	0.860028	1.31 (<.0001)	**1.43 (<.0001)**	**1.55 (<0.0001)**

Table 12: Coefficients for news score 5 in model 5, both with and without lagged volume

company	alpha	beta	volume volume coefficient($\times 10^{-8}$)	News score in model 5 with volume($\times 10^{-11}$)	News score in model 5 without volume($\times 10^{-11}$)
Alcatel	0.052917	0.940166	4.777 (<.0001)	**2.22 (<.0001)**	**3.3(<0.0001)**
Bayer	0.099995	0.872391	1.798 (<.0001)	-1.06 (0.9391)	-5.5(0.7015)
BBVA	0.07025	0.91003	1.15 (<.0001)	-5.42 (0.7477)	8.25(0.6372)
Deutsche Telecom	0.05936	0.905872	1.208 (<.0001)	**0.557 (0.0005)**	**0.511 (0.004)**
Deutsche Borse	0.164643	0.818504	2.052 (<.0001)	**616.9 (<.0001)**	**647.8 (<0.0001)**
ENEL	0.03665	0.943788	1.673 (<.0001)	**0.28 (0.0035)**	**0.444 (0.0003)**
ENI	0.045893	0.939713	0.9881 (<.0001)	0.0696 (0.3302)	0.078 (0.3209)
France_Telecom	0.070743	0.914509	2.251 (<.0001)	-0.143 (0.6732)	0.119 (0.7551)
Generali Asset	0.067742	0.911944	0.8058 (<.0001)	-6.98 (0.2017)	-4.82 (0.428)
Iberdrola	0.110681	0.843288	2.894 (<.0001)	0.173 (0.8484)	-95.7
Intesa Sanpaolo					**60.8 (0.0095)**
Muench Rueckvers	0.11248	0.87475	0.9912 (<.0001)	-2.86 (0.5442)	-7.14(0.1508)
Nokia	0.038367	0.953422	1.06 (<.0001)	**34.3 (<.0001)**	**37.4(<0.0001)**
Repsol	0.10506	0.870056	1.279 (<.0001)	6.87 (0.3966)	5.99 (0.4538)
RWE	0.099732	0.880982	1.333 (<.0001)	**134.2 (<.0001)**	**147 (<0.0001)**
Santander	0.075611	0.909552	1.142 (<.0001)	-1.98 (0.4402)	-2.28 (0.359)
SAP	0.09402	0.883843	1.102 (<.0001)	**0.619 (0.0010)**	**14.3 (<0.0001)**
Siemens	0.098831	0.885354	0.8219 (<.0001)	**-0.717 (<.0001)**	**-0.834 (<0.0001)**
Telefonica	0.071473	0.905319	1.1(<0.0001)	**0.183(0.001)**	**0.343 (0.0212)**
Vow	0.120008	0.859543	1.303 (<.0001)	**0.907 (<.0001)**	**1.03 (<0.0001)**

7 Appendix 2 : Impact function

Company	Time to first crest(min)	Time to second peak(min)	time lag between crest and peak(min)
Deutsche Bank	76	98	22
Sap	87	160	73
Deutsche Telecom	80	120	40
Volkswagen	53	72	19
Eni	80	156	76
Iberdrola	60	131	71
Telefonica	55	69	14
Alcatel-Lucent	44	79	35
France Telecom	100	160	60
Enel	61	114	53
Bayer	84	124	40
Bbva	102	141	39
Deutsche Boerse	68	112	44
Generali Asset	95	155	60
Intesa sanpaolo	103	147	44
Muench Rueckvers	17	83	66
Nokia	107	180	73
Repsol	112	122	10
Rwe	69	111	42
MEAN	76.47	122.84	46.36
SD	24.58	32.57	20.69
TRIMMED MEAN 10%	77.88	122.64	46.76
MEDIAN	80	122	44

Table 13: Summary characteristics of the impact curves

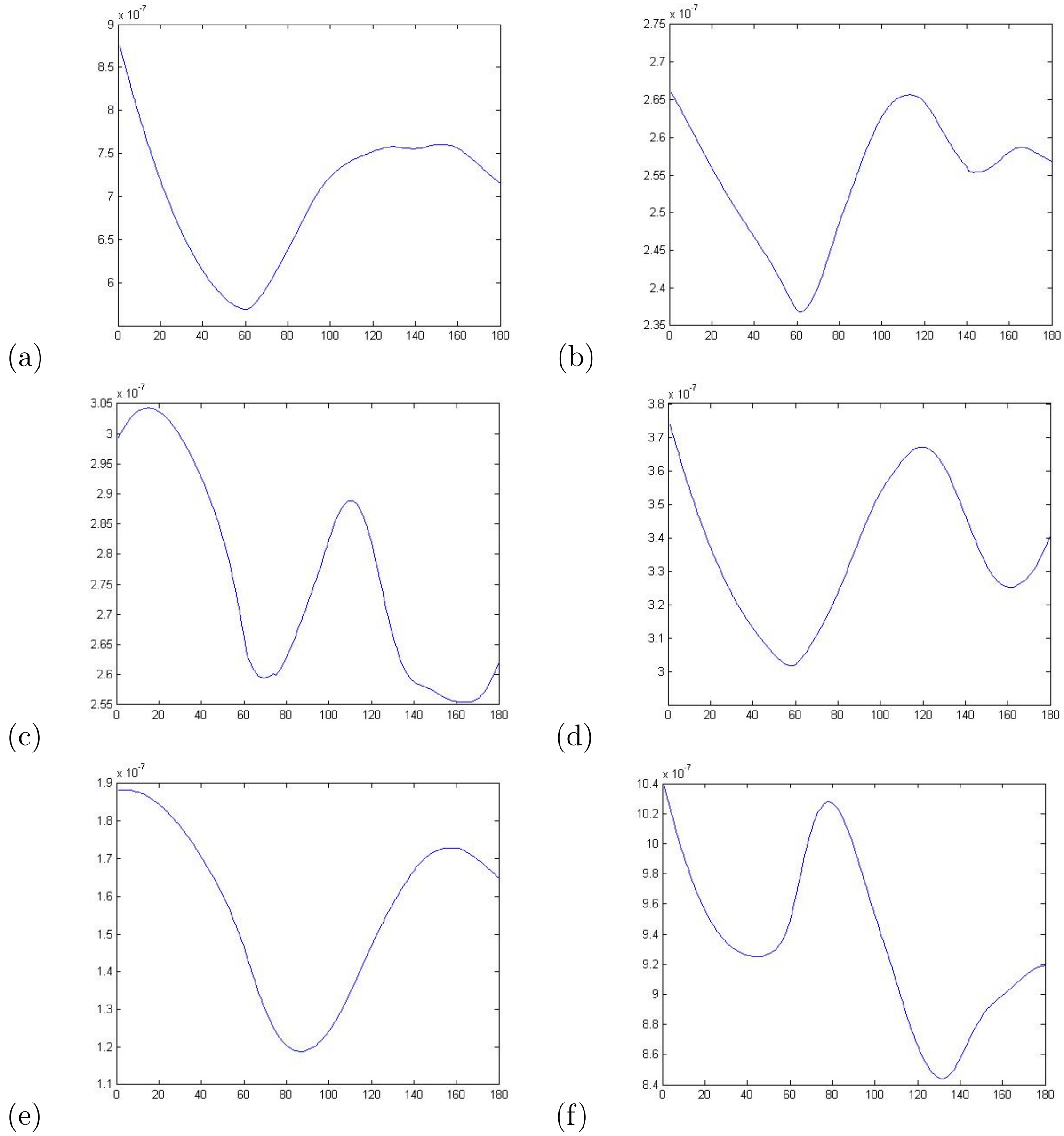
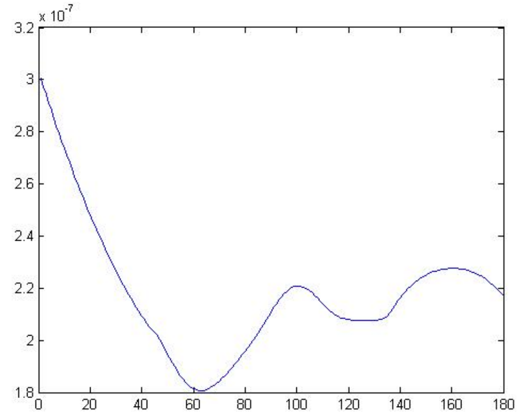
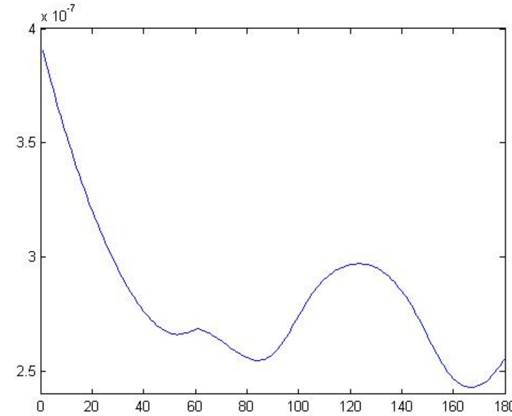


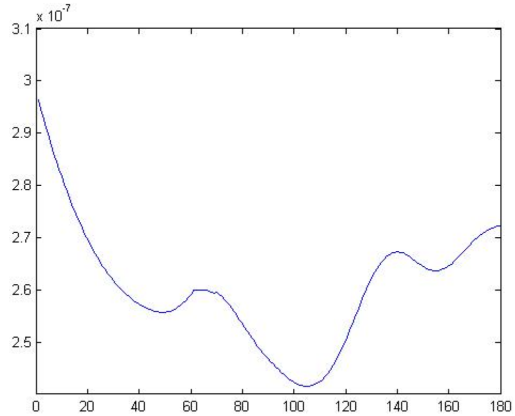
Figure 7.1: Plots of empirical impact functions (a) Iberdrola (b) Enel (c) Rwe (d) Daimler (e) Sap (f) Alcatel-lucent



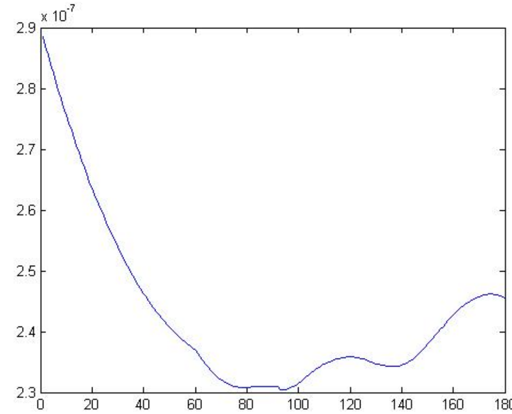
(g)



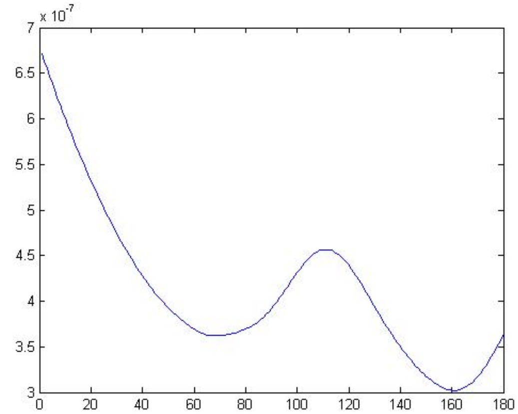
(h)



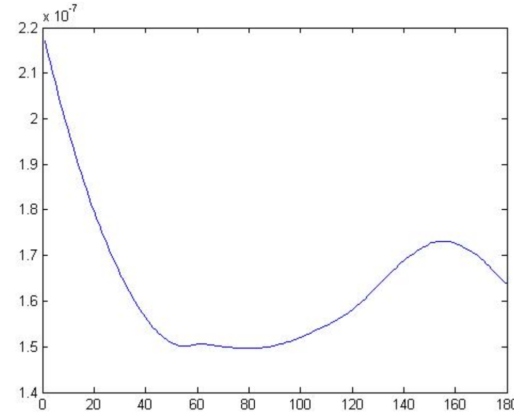
(i)



(j)

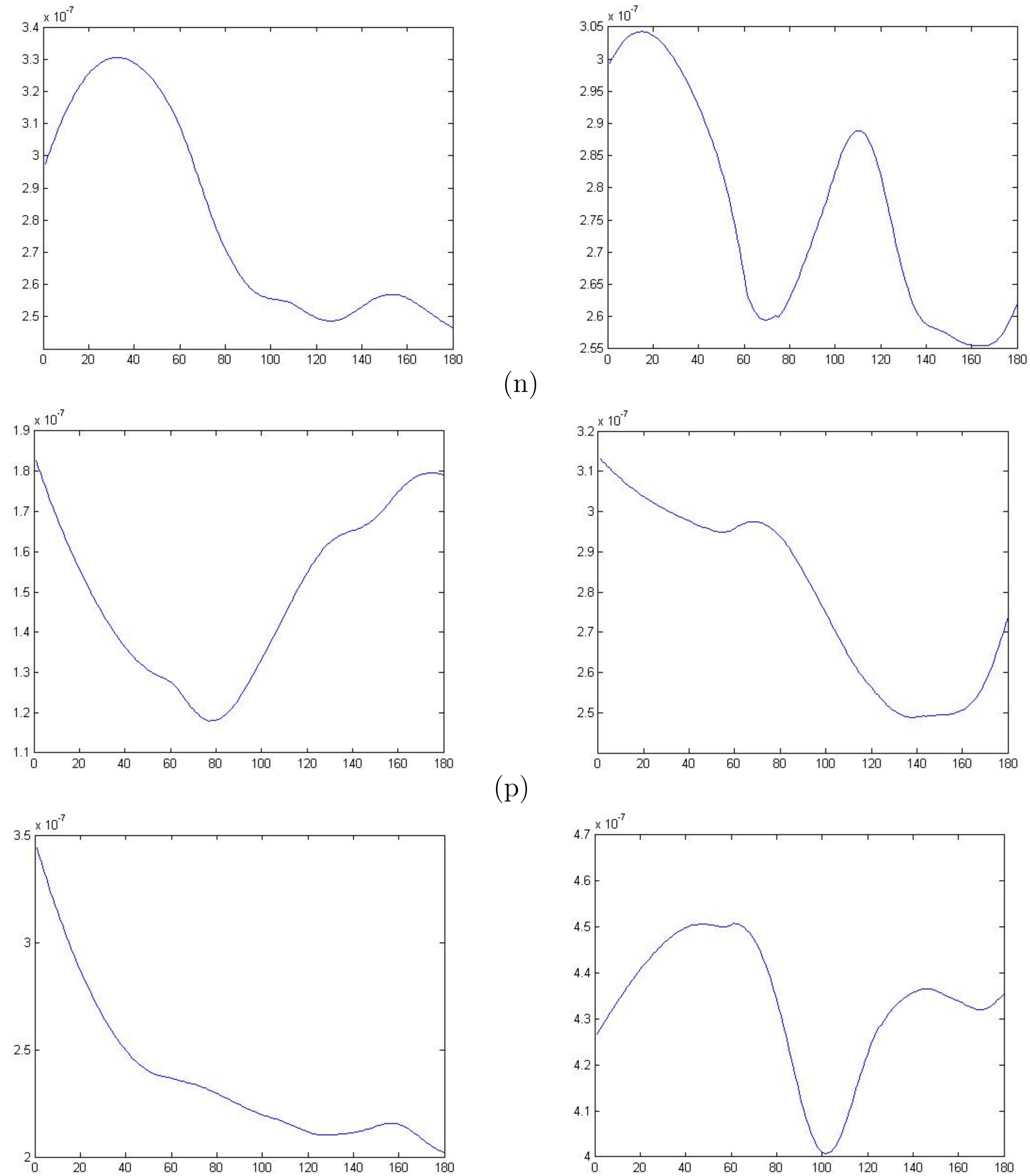


(k)



(l)

Figure 7.2: Plots of Empirical Impact Functions (g) Allianz (h) Bayer (i) BBVA (j) Deutsche telecom (k) Deutsche Borse (l) Eni

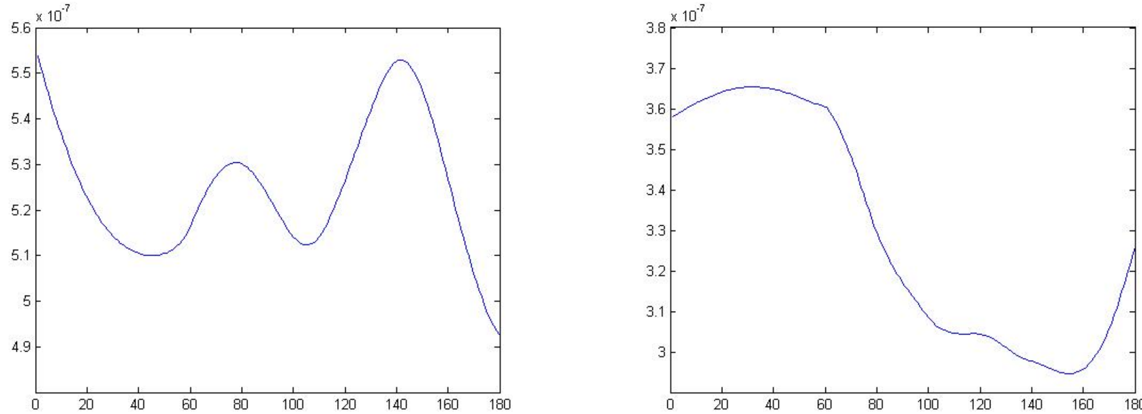


(m) (n)

(o) (p)

(q) (r)

Figure 7.3: (m) generali asset (n) Rwe (o) Siemens (p) telefonica (q) Vow (r) intesa sanpaolo



(s) (t)

Figure 7.4: (s) santandar (t) repsol

8 References

Akgiray, V. (1989) Conditional heteroskedasticity in time series of stock returns: evidence and forecasts, Journal of Business, 62(1), 55-80.

Andersen, T.G., Bollerslev, T, Diebold, F.X., and Labys, P. (2001) Exchange rate returns standardized by realized volatility are (nearly) Gaussian, Multinational Finance Journal, 4, 159-179.

Andersen, T.G., Bollerslev, T, Diebold, F.X., and Labys, P. (2001) The distribution of realized exchange rate volatility, Journal of the American Statistical Association, 96 (453), 42-55.

Andersen, T.G., Bollerslev, T., Diebold, F.X., and Labys, P. (2003) Modeling and forecasting realized volatility, Econometrica, 71, 529-626.

Barber, Brad. M., and Odean Terrance (2008) All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors, The Review of Financial Studies, 21(2), 785-818.

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

Black, F. (1976) Studies in stock price volatility changes, Proceedings of the 1976 Business Meeting of the Business and Economic Statistics Section, American Statistical Association, 177-181.

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

Bollerslev, T., Chou, R.Y., and Kroner, K.F. (1992) ARCH modeling in finance: a selective review of the theory and empirical evidence, Journal of Econometrics, 52, 5-59.

Brooks, C. (1998) Predicting stock index volatility: can market volume help? Journal of Forecasting, 17, 59-80.

Brooks, C., and Persand, G. (2003) Volatility Forecasting for Risk Management, Journal of Forecasting, 22, 1-22.

Christie, A.A. (1982) The Stochastic Behavior of Common Stock Variances: Value, Leverage and Interest Rate Effects, Journal of Financial Economics, 10, 407-432.

Clark, P. (1973) A subordinated stochastic process model with finite variance for speculative prices, Econometrica, 41, 135-155.

Cleveland, W.S., Devlin, S.J. (1988). Locally-Weighted Regression: An Approach to Regression Analysis by Local Fitting. Journal of the American Statistical Association 83 (403): 596–610.

Corsi, F., Zumbach, G., Muller, U.A., and Dacorogna, M.(2001) Consistent high precision volatility from high-frequency data, Economic Notes, 30(2), 183-204.

Engle, R.F. (1982) Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. inflation, Econometrica, 50, 987-1008.

Engle, R F. And Ng, V. (1993) Measuring and testing the impact of news on volatility, Journal of Finance, 48, 1749-1778.

Epps, J., and Epps, M. (1976) The stochastic dependence of security price changes and transaction volumes: implications for the Mixture-of-Distributions hypothesis, Econometrica, 44, 305-321.

Glosten, L.R., Jagannathan, R., and Runkle, D.E. (1993) On the relation between the expected value and the volatility of the nominal excess return on stocks, Journal of Finance, 48(5), 1779-1801.

Jones, C. Kaul, G. and Lipson M. (1994) "Transactions, Volume and Volatility" Review of Financial Studies, 7, 631-651.

Kalev P.S, Liu, W.M,Pham P.K,Jarnecic E (2004) Public information arrival and volatility of stock returns, Journal of Banking & Finance, 28,1141-1467 .

Lamoureux C.G., Lastrapes W.D., (1990) Heteroscedasticity in stock return data; volume versus GARCH effects. Journal of Finance, 45, 221-229

Lamoureux C.G., Lastrapes W.D., (1994) Persistence in variance,structural change and the GARCH model. Journal of Bussiness and economic statistics, vol 8, No 2.

Leinweber, Davis. and Sisk Jacob (2011), Relating news analytics to stock returns, in The Handbook of News Analytics in Finance, edited by L.Mitra and G.Mitra, Wiley Finance, 149-172

Mandelbrot, B. (1963) The variation of certain speculative prices, Journal of Business, 36(3), 394-419.

McAleer, Michael, and Medeiros, Mercelo C. (2008) Realized volatility: A review, Econometric Reviews, 27:1, 10-45.

Merton, R.C. (1980) On estimating the expected return on the market: An exploratory investigation, Journal of Financial Economics, 8, 323-361.

Moniz, A., Brar, G., Davies, C. And Strudwick, A. (2011) The impact of news flow on asset returns: An empirical study, in The Handbook of News Analytics in Finance, edited by L.Mitra and G.Mitra, Wiley Finance, 211-230

Nelson, D.B. (1991) Conditional heteroskedasticity in asset returns: A new approach Econometrica, 53, 347-370. Shalen, Katherine, T. (1993) Volume, volatility and the Dispersion of Beliefs, Review of Financial Studies, 6, 405-434.

Sharma, Jandhyala, L., Mougoue, Mbodja, and Kamath, Ravindra. (1996) Heteroscedasticity in stock market indicator return data: volume versus GARCH effects, Applied Financial Economics, 6, 337-342.

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

West, K.D., Edison H.J., Cho, D. (1993) A utility-based comparison of some models of exchange rate volatility, Journal of International Economics, 35, 25-45.

Zhou, Ben. (1996) High –frequency data and volatility in foreign-exchange rates, Journal of Business & Economic Statistics, 14(1), 45-52.