News Intensity and Conditional Volatility

Jean-Gabriel Cousin[a], Tanguy de Launois[b*]

[a]*ESA, Université de Lille II, France*

[b]*IAG, Université catholique de Louvain, Belgium*

April 2005

ABSTRACT

The relation between information flow and asset prices behavior is one of the key issues of modern finance. Our study investigates more closely the link between frequency of information arrivals and stock return volatility. It aims precisely to test empirically the mixture of distribution hypothesis and to check whether the stock returns distribution is driven by the frequencies of information arrivals on the Paris stock Exchange (Euronext). We analyze the impact of news on volatility at the firm-level. We opt for a model with two (Markov switching) regimes of volatility that we apply to all stocks pertaining to the CAC40 index from January 1999 to December 2003. We find a positive and significant but marginally decreasing impact of the daily frequency of information arrivals on the probability to be in a state of high volatility for each of the 40 companies considered. The subsequent model for panel data allows us to conclude that this impact crucially depends on the timing and the subject of the news release. Asymmetry and informational content issues are also investigated. Results are consistent with previous literature, although we show that any asymmetric effect disappears once the news informational content is accounted for.

JEL codes: G14
EFA Classification Code: ME

Keywords: News, (conditional) Volatility, Public information arrivals, Mixture of distribution hypothesis, Communication intensity, Markov Switching Regression Model.

---

* Corresponding author:
1, Place des Doyens 1
1348 Louvain-la-Neuve
Belgium
e-mail address: delaunois@fin.ucl.ac.be

Other author :
1 Place Déliot
59000 Lille
France
e-mail address: jgcousin@univ-lille2.fr

We are grateful to Jacques Chardon and Factiva for having offered access to their data base. We are grateful to participants of the AFFI 2004 (Paris – December) for their constructive comments. We especially wish to thank E. de Bodt, M. Levasseur and N. Aktas for their advice. All errors remain ours.

1. Introduction

The relation between information flow and asset prices behavior is one of the key issues of modern finance. How, why and under which conditions do financial news impact stock price volatility count among the typical (and still debatable) questions in this research area. The clear knowledge and the good understanding of such important issues are of first importance for listed companies. The behavior of stock prices around profit warning announcements is a typical (and probably extreme) example of these phenomena. Serge Tchuruk, Alcatel's CEO, still probably remembers the disastrous consequences of his injudicious profit warning of September 17, 1998. At first sight, the news release looked quite innocent: "The operational result of the telecommunications sector (…) will suffer from the large decrease in investments (…) and from the crisis in South-East Asia and in Russia. Alcatel reckons that the operational performance in 1998 will not reach the expected level." Among other things, Alcatel said it expected to make a profit in 1998 of €1.1 billion, down sharply from an earlier forecast of €1.4 billion. The market reaction was terrible: shares of the company lost 38% in one day, involving that not less than €10 billion of Alcatel's market capitalization vanished in a couple of hours. Never had the "Paris bourse" experienced such a dramatic stock crash.

That kind of news hardly makes only one victim. Sector spill over effects of such announcements have been quite common. On September 16, 2004, after Celestica, a large-cap electronics manufacturer, cut its outlook for the third quarter, the whole sector of makers of electric parts and equipment as well as providers of related services dropped: Power-One tumbled 6.4%, Plexus fell 5.4%, Littlefuse shed 3.3% and Benchmark Electronics lost 10% on the New York Stock Exchange. The communications-technology sector was affected the same way, since Stratex Networks dropped 4.4%, Mindspeed Technologies lost 3.3% and MRV Communications fell

3.2%. Financial communication may thus strongly impact stock return, volume and volatility of any company. Such a point deserves to be studied and analyzed carefully.

Our primary objective is to investigate the relation between public information and stock volatility. To what extent does the discontinuous frequency of news releases affect the volatility of the daily stock returns? And what are the factors strengthening or weakening this relation? These are precisely the main questions that we aim to answer in this paper. They are crucial from both a practical and an academic perspectives, since recent academic works have put into light the importance of developing a better understanding of the determinants of return volatility. Easley and O'Hara (2001) bring some evidence that the less transparent and disclosing a company is, the higher its cost of capital will be. Such a conclusion already offers some insight of how significant the influence of news releases (in quality and in quantity) on stock prices can be. Campbell and al. (2001) find that the proportion of firm-specific variance in the total stock variance has steadily been increasing for the last thirty years and thus provides us with another good reason to investigate the relation between firm-specific news and idiosyncratic volatility. As a matter of fact, one may wonder whether the regular increase in idiosyncratic variance detected by Campbell and al. (2001) could possibly be linked with the obvious contemporaneous increase in the financial communication intensity. Our results seem to support such an assertion. Eventually, Goyal and Santa Clara (2002) detects a positive relation between idiosyncratic variance and expected stock return. For idiosyncratic risk is not perfectly diversifiable and investors are assumed risk-averse, an increase in the firm-specific part of the stock volatility can induce an increase in the cost of capital[1].

Since we focus on the stock volatility, an adequate modelisation of its dynamic is obviously of prime importance. In the literature, its time-varying nature is a well-known empirical fact. Presence of conditional heteroscedasticity is well documented in most

[1] However, this point remains controversial. In obvious contradiction with Goyal and Santa Clara (2002), Wei and Zhang (2004) claim that idiosyncratic risk does not matter.

financial time-series. As a natural consequence, several authors have tried to account for this volatility persistence effect in alternative econometric frameworks like ARCH (see Engle (1982)) and GARCH[2] (see Bollerslev (1988)) approaches. By expressing the stock idiosyncratic variance at time t as a linear function of the past stock idiosyncratic variances (at *t-1*, *t-2*, etc.) and innovations, these models provide conditional estimates of the stock volatility while taking the persistence volatility effect into account.

Markov switching regression models (MSR) appear as another interesting way to capture conditional heteroscedasticity evolutions. In such a non-linear dynamic framework, stock returns distribution is supposed to switch from a low to a high volatility regime (and vice-versa) according to some fixed transition probabilities. MSR models exhibit a more general structure than the GARCH models, since they impose fewer assumptions on the functional form of the conditional volatility. Moreover, GARCH models may sometimes present some annoying convergence problems, while MSR models are generally more reliable in this respect. Eventually, GARCH models do not rest on a clear theoretical justification (they are sort of reduced form models), whereas MSR models are implicitly based upon such a powerful argument, which is known as the Mixture of Distribution Hypothesis (MDH). Suggested for the first time by Clark (1973), it posits that the joint distribution of daily return and volume can be modeled as a mixture of bivariate normal distributions. Specifically, they are contemporaneously dependent on an underlying mixing variable that represents the flow of information. As a consequence, the variance of returns at a given interval is expected to be proportional to the rate of information arrival at the market. About fifteen years later, Roll (1988) shows empirically that the sample variance and the kurtosis for example can reveal something about the probability of information and the difference between the information-related distribution and the non-information-related distribution of returns, which is totally consistent with the MDH. One year later, Ross

[2] ARCH and GARCH stand respectively for autoregressive conditional heteroscedasticity and generalized autoregressive conditional heteroscedasticity (models).

(1989) brings some convincing theoretical evidence supporting the MDH and demonstrates that in an arbitrage-free economy, the volatility of prices is directly related to the rate of flow of information to the market. An avowed objective of our study is precisely to test empirically that mixture of distribution hypothesis and to check whether the stock returns distribution is driven by the frequency of information arrivals on the Paris stock Exchange (Euronext).

Though GARCH and MSR models give useful and quite reliable estimates of the conditional stock variance, they still remain endogenous models not dedicated to explain observed changes in stock volatility and to point out the core factors driving its behavior. In our quest for an innovative explicative model of the stock volatility, the MDH is of prime importance, since it argues that the stock volatility is directly related to the information frequency. Therefore it largely justifies the introduction of the "daily number of news releases on a stock" in the basic equation of our intended explicative model as the most appropriate explanatory variable.

If many authors tried to establish the reality of the positive relation between flows of information and stock (or market) volatility, the empirical evidence of such a positive relation is everything but overwhelming, probably because of the difficulty to find good empirical proxy of information arrivals and as a result of the poor volatility estimates that were traditionally used. The difficulty to precisely measure information arrivals appears in the variety of proxies employed by previous empirical studies on the topic. Berry and Howe (1993) use the number of daily newspaper headlines and earnings announcements and Ederington and Lee (1993) investigate the importance of macroeconomic news, whereas Mitchell and Mulherin (1994) employ the number of specific stock market announcements in order to test the impact of the rate of information on the market volatility. « However, the use of unconditional volatility measures such as absolute daily market returns in these studies often generates weak or inconclusive results regarding the news–volatility relation »[3]. Indeed, such a proxy is known to be quite noisy, and the presence of conditional heteroscedasticity in the returns time-series may significantly alter the quality of the results. It is precisely to account for this well-known phenomenon of conditional heteroscedasticity that Kalev and al. (2004) choose to test the relation between firm-specific announcements (as a proxy for information flows) and volatility on the Australian Stock Exchange in a GARCH framework. Their analysis reveals a positive and significant impact of the arrival rate of the selected news variable on the conditional variance of stock returns, even after controlling for the potential effects of trading volume and high opening volatility. They find the respective coefficients of these two last control variables to be positive and significant, except on a daily basis. Eventually, they split all their press releases into different categories according to their subject, since Andersen (1996) argues that different types of news have different stochastic arrival processes, so that we may legitimately expect them to have a different impact on the conditional stock volatility.

Our work is in the line of the above mentioned publications. Its contributions are of three kinds: they concern the data used as well as the methodology adopted and the results obtained.

In order to test empirically whether information flows drive stock return volatility, we apply our econometrical model to all stocks pertaining to the CAC40 index from January 1, 1999, to December 31, 2003. Datastream provides us with the series of daily stock prices, while Factiva gives us access to a particularly rich and advantageous news sample composed of all time-stamped societal and industrial Reuters news releases concerning all companies selected above. This kind of financial communication counts undoubtedly among the freshest and most relevant public information that investors can collect when trading on the market. Furthermore, since we are only interested in the

[3] Kalev et al. (2004), p. 1442

impact of information intensity on the idiosyncratic stock variance, we restrict our investigation to firm-specific and industrial-specific news and chose not to introduce any market or macroeconomic news in the sample (as it will be explained below, we will control for possible market and macroeconomic shocks on stock variance through controlling for the market variance in the regression equation). Finally, for we want to test the link between the daily number of firm-specific news releases and the corresponding stock volatility, it is necessary to know which company (or companies) is (or are) concerned by each news release. Our sample provides us with such an interesting item, among many others. It allows us to refine the classical aggregate market-level analysis into a more subtle and accurate firm by firm study.

Our methodology differs in many respects from the above mentioned previous studies on the topic. We analyze the impact of news on volatility at the firm-level, i.e. company by company, and not at the market-level, i.e. impact of news on the index. It allows us to capture the entire "firm-specific effect" contained in the news which is obviously lost at the aggregate level. We also employ conditional volatility measures (instead of unconditional volatility measures) in order to take the volatility persistence effect into account. We first replicate on a daily basis the GARCH approach of Kalev and al. (2004) with our specific data set at the firm-level. We then opt for a market model with two (Markov switching) regimes of volatility, which is more robust and general than the GARCH model used by Kalev and al. (2004) because it needs less strong assumptions on the conditional volatility process. Moreover, such a MSR model allows us to explicitly take advantage of the central Roll (1988)'s idea[4] of a mixed return distribution ("with news" and "no news") driven by news arrivals and to test the relevance of such an intuition. The subsequent econometric methodology consists basically of two steps: estimating the daily probabilities of being in a high volatility state thanks to the developed two-state market model and regressing these obtained

[4] The MDH in fact.

probabilities on some possibly explicative variables (like the daily number of news) in order to understand the determinants driving the stock returns volatility. We also perform a panel analysis to better judge of the overall significance and robustness of our results. Like Kalev and al. (2004), we control for the daily market volume of trades though we find it to be insignificant. Unlike Kalev and al. (2004), we control for factors such as the market conditional volatility and the sector conditional volatility, which turn out to be highly significant. We also choose to perform the same regressions on different news subsets selected on the basis of various criteria, like the subject or the timing of the news release. In a last and deeper analysis, we investigate the asymmetry issue and we propose a measure of the information content of the news releases.

Our main finding is that public information flows strongly affect the stock returns distribution. In the GARCH framework, our results are consistent with those of Kalev and al. (2004). The news frequency seems to be largely responsible for the persistence volatility effect, and its impact on the volatility level is substantial. In the MSR model framework, we find a positive and significant impact of the daily frequency of information arrivals on the probability to be in a state of high volatility for each of the 40 companies considered. The subsequent model for panel data confirms this result and demonstrates that this impact is marginally decreasing. The regressions on various data subsets offer very interesting results: news released during trading hours seems to have a much stronger impact than news released during non-trading hours. In terms of subject, press releases concerning *regulation and government policy*, *earnings projections* (especially the subset of *analyst comments and recommendations*) and *mergers and acquisitions* (in order of decreasing importance) appear to have the greatest impact on the stock volatility. Eventually, the investigation of the asymmetry issue and the development of a news informational content proxy lead us to conclude that one good and one bad announcements with identical informational content affect the conditional volatility the same way, so that the classical asymmetry effect reported by the literature simply disappears.

Our paper is organized as follows. Section 2 presents the GARCH and the Markov switching regimes model. Section 3 describes the used data sets: stock prices and news. Section 4 displays and comments results. Section 5 concludes and provides some perspectives for future research.

2. Model specification

One of the core intermediate objectives of our work is to obtain adequate estimates of the daily stock volatility. In this aim, we choose to apply two different frameworks: GARCH model, as used in Kalev and al. (2004), and Markov switching model that is thought to be more relevant and appropriate than the former.

The GARCH model

First of all, we propose to simply replicate the study of Kalev and al. (2004) on our extensive data set. Their article aimed at testing the relation between firm-specific announcements (as a proxy for information flows) and volatility on the Australian Stock Exchange in a GARCH framework. By testing the same relation in the same framework on a richer and more extensive dataset (namely Euronext news), we want to check that we get results consistent with those of Kalev and al. (2004). However, the study we intend to perform is different from theirs in two respects: while Kalev and al. (2004) investigate the volatility behavior on an intraday and daily scale and were mainly interested in the market volatility, we choose to focus on a daily scale on the individual stock volatility.

GARCH model have now become quite common in the modelisation of uncertainty in financial asset returns. We will use it as follows:

$$R_{i,t} = \alpha_i^{MM} + \beta_i^{MM} R_{m,t} + \varepsilon_{i,t} \text{ where } \varepsilon_{i,t} \left| \Omega_{t-1} \sim N(0, \sigma_{i,t}^2) \right. \quad (1)$$

$$\sigma_{i,t}^2 = \omega_i + \alpha_i \varepsilon_{i,t-1}^2 + \beta_i \sigma_{i,t-1}^2 \quad (2)$$

where: $R_{i,t}$ is the return of stock i at interval t;

$R_{m,t}$ is the return of market index at interval t;

α_i^{MM} and β_i^{MM} are the market model parameters for the stock i;

$\varepsilon_{i,t}$ are the serially uncorrelated errors of stock returns i with mean zero;

and $\sigma_{i,t}^2$ is the conditional variance of $\varepsilon_{i,t}$.

The coefficients α_i and β_i reflect the dependence of the current volatility of stock i upon its pasts levels and the sum $\alpha_i + \beta_i$ indicates the degree of volatility persistence.

According to Lamoureux and Lastrapes (1990) and Kalev and al. (2004)'s results, the time-varying pattern of conditional volatility may be generated by serial correlation in the information arrival process. This argument implies that when a proxy for information flows is inserted directly into the conditional variance equation, most of the observed volatility persistence is expected to disappear. Like Kalev and al (2004), we adopt as a proxy the number of all firm-specific news events announced to the stock market per interval t (N_t), but contrary to them we run the analysis at the individual firm level ($N_{i,t}$) rather than at the market level in order to better and more precisely capture the impact of news announcements on the stock volatility. Importantly, $N_{i,t}$ cannot account for the flow of private information, it does just encompass a major component of the public information set that drives stock prices and volatility. In our first analysis $N_{i,t}$ is set as an exogenous variable in the conditional variance equation of the GARCH (1,1) specification as follows:

$$\sigma_{i,t}^2 = \omega_i + \alpha_i \varepsilon_{i,t-1}^2 + \beta_i \sigma_{i,t-1}^2 + \lambda_i N_{i,t} \quad (3)$$

Two important results follow from the estimation of the above equation. Firstly, the significance of the coefficient λ_i provides evidence about whether the rate of public information arrival influences volatility of the stock i in the presence of conditional heteroscedasticity. Secondly, if the presence of GARCH effect is largely induced by the nature of the information flows, we may legitimately anticipate a significant decrease in the persistence of volatility, $\alpha_i + \beta_i$. Since $N_{i,t}$ does not cover all sources of information, we do not expect volatility persistence to disappear entirely after the inclusion of $N_{i,t}$.

After this brief work of "replication", we will continue our quest for an innovative explicative model of the stock volatility. In that aim, the MDH is of prime importance, since it argues that the stock volatility is directly related to the information frequency. Therefore it largely justifies the introduction of the "daily number of news releases on a stock" in the basic equation of our intended explicative model as the most appropriate explanatory variable. If our first results, based on the GARCH framework, are in line with the standard MDH, which suggests that return volatility is proportional to the rate of information arrival, we will be encouraged to go one step further and to think of a model that could overcome the drawbacks of the classical GARCH models.

The two-state market model (TSMM)

We propose a new way to model the stock's variance behavior that is perfectly in line with the MDH and Roll's (1987) results. According to Roll, the true return generating process seems to be better described by a mixture of two distributions: one corresponding to a state of information arrival, and the other to the normal return behavior. In order to study the impact of news on volatility at the firm level, our new approach is based on a combination of the well-established market model (Sharpe, 1963) and the more recent Markov Switching Regression models (MSR), largely introduced and developed by Hamilton (1989 and 1994), and significantly extended in Krolzig (1997). Our initial intuition is simple: if the occurrence of firm-specific events may have a significant impact on the variance of the firm's return generating process, we must capture this perturbation by using a regime switching model.

In this framework, we assume that the error $\varepsilon_{i,t}$ is state dependent. S_t denotes our state variable. We consider the case of a two-state regime model (mixture of two distributions with different variances). More precisely, we have a low-variance regime ($S_t = 1$) and a high-variance regime ($S_t = 2$):

$$\begin{aligned} y &= X\delta + \varepsilon_1 \quad if\, S_t = 1 \\ y &= X\delta + \varepsilon_2 \quad if\, S_t = 2 \end{aligned} . \tag{4}$$

The variance of the residuals for each state is given by:

$$\begin{aligned} E[\varepsilon_1\varepsilon_1'|X] &= \sigma_1^2 I \quad if\, S_t = 1 \\ E[\varepsilon_2\varepsilon_2'|X] &= \sigma_2^2 I \quad if\, S_t = 2 \end{aligned} , \tag{5}$$

where $\sigma_2^2 > \sigma_1^2$.

We assume that the transition between the two regimes is governed by a Markov chain of order 1, for which the transition matrix is given by:

$$P = \begin{bmatrix} p_{11} & 1-p_{22} \\ 1-p_{11} & p_{22} \end{bmatrix}, \tag{6}$$

where $p_{ij} = p(S_t = i \mid S_{t-1} = j)$ corresponds to the probability of going from state j to state i. The unconditional probability of the regime is given by (Hamilton, 1994, p. 683):

$$\begin{aligned} p(S_t = 1) &= \frac{1-p_{22}}{2-p_{11}-p_{22}} \\ p(S_t = 2) &= \frac{1-p_{11}}{2-p_{11}-p_{22}} \end{aligned} . \tag{7}$$

The intuition underpinning our choice is simple: we anticipate that firm-specific events will impact the return variance. It could be argued that it is in contradiction with the semi-strong form of efficiency hypothesis, under which stock prices should adjust immediately to any public information announcement. It is in fact not the case if we take into account the uncertainty attached to firm-specific events. We will suppose that the return generating process can be adequately modeled using a two-regime process[5], one

[5] This hypothesis is supported by unreported results. Using the approach developed in Krolzig (1997), we find that a two-regime model is an adequate representation of the return generating process in the vast majority of cases. Three-regime decomposition appears to be justified only in the presence of strong outliers.

regime with normal variance and one regime with high variance (firm-specific event regime). Note that, in both regimes, the MM (market model) parameters are assumed to be the same. Therefore the return generating process is:

$$\begin{aligned} &R_{i,t} = \alpha_i + \beta_i R_{m,t} + \varepsilon_{i,1,t} \quad \text{if } \mathrm{S_t} = 1 \text{ with } \varepsilon_{i,1,t} \sim \mathrm{N}(0, \sigma_{\mathrm{i},1}) \\ &R_{i,t} = \alpha_j + \beta_i R_{m,t} + \varepsilon_{i,2,t} \quad \text{if } \mathrm{S_t} = 2 \text{ with } \varepsilon_{i,2,t} \sim \mathrm{N}(0, \sigma_{\mathrm{i},2}) \\ &\text{and} \quad \sigma_{\mathrm{i},2} > \sigma_{\mathrm{i},1} \end{aligned} \tag{8}$$

where S_t is a state variable taking the value "1" if we are in the low variance regime and the value "2" if we are in the high variance regime. In the sequel, we will refer to equation (8) as the two-state market model (TSMM). The proposed model is a direct and parsimonious extension of the classical MM. As the regime state variable S_t is not directly observable, we have to specify its statistical properties. The model we propose is therefore based on the estimation of six parameters (α, β, σ_1, σ_2, p_{11} and p_{22}) and, while much more flexible than the classical MM model, remains parsimonious.

The estimation of Markov Switching Regression models is fully presented in Hamilton (1994). It is based on a maximum likelihood approach, for which an efficient estimation algorithm has been developed. The estimated probability of being in a specific state at a specific date is one of the interesting by-products of the advocated approach. It allows one to look for the reasons explaining an increase in the variance of stock returns; in our case, to look for the link between information flow and conditional variance.

The use of this Markov Switching Regression model provides numerous interesting features. Beyond the above-quoted possibility of obtaining the estimated probability of being in a specific regime at a specific date, this specification is in line with Roll's (1987) intuition. In his Presidential Address to the American Finance Association, he clearly highlighted that the true return-generating process seems to be better described by a mixture of two distributions, the first one corresponding to a state of information arrival and the other to the normal return behavior: it's clearly a good candidate for testing the mixture of distribution hypothesis. Once we have estimated the probability for a firm of being in a high variance regime at a specific date, we can test whether it is linked with information flow.

The specification has other attractive features:

- the Gaussian conditional distribution of returns could be misleading. While the model imposes a Gaussian assumption for the return distribution in each state, as shown in Hamilton (1994) and Krolzig (1997), it allows to capture skewness and kurtosis in the unconditional distribution;
- the Markov Switching Regression framework also takes conditional heteroscedasticity into account without imposing a specific form on the conditional dependence of the variance (as in the (G)ARCH framework);
- finally, the estimation process provides us with the estimated variance in each regime, the probability of being in a specific regime at a specific date, and the estimated transition probabilities[6].
- in a certain manner, the TSMM, once we have estimated the probability for a firm of being in a high variance regime at a specific date, could be a better framework than the GARCH model (via the estimated λ_i) in order to estimate the impact of the public information rate on the stock volatility at the individual firm level. This may also give some additional robustness to the news-volatility relation in the pattern of the daily stock return volatility.

3. Sample selection

Our firms' universe is composed of the CAC40 index stocks. It provides us with a broad and representative sample of the French Stock Market since it accounts for not less than 70% of the French stock market value. The total market value of the index on

[6] All estimations presented in this paper have been realized under the Ox econometric software, using the Krolzig MSVAR package. We thank Professor J. Hamilton for advising the use of this package. It is freely downloadbale at: http://www.nuff.ox.ac.uk/Users/Doornik/index.html.

November 11, 2003 was €722 billion. The mean (median) company market value was €18 billion (€13 billion). For all firms included in this index, we use daily prices from January 1, 1999 to December 31, 2003. The three largest sectors in the index were the financials (22.29%), resources (oil) (18.27 %) and non cyclical consumers' goods (11.82%). We use as a market portfolio the CAC40 index. The data are obtained from Datastream, accessed at the university of Lille 2.

Our news sample is provided by Factiva. It is composed of all corporate and industrial news concerning firms pertaining to the CAC40 index and recorded by Reuters from January 1, 1999 to December 31, 2003. For each news wire, we have got the following fields: the accession number (AN), headline (HD), word count (WC), publication date (PD) and time (ET), source name (SN) and code descriptor (SC), lead paragraph (LP), company (concerned by the news release) code (CO), industry code (IN), subject code (NS), region code (RE), Dow Jones codes (DJIC) and descriptors (DJID), information provider codes (IPC) and Reuters codes (RBBCM). One piece of news can of course concern several companies, industries and subjects.

To avoid any redundancy and duplicate announcements that do not bring any additional information value, we restrict the sample to news released by Reuters only. For the same reason, we eliminate all news releases with the same headlines and lead paragraphs. Eventually, after having excluded news items with one missing field or more, we are left with 76341 financial communications over the whole five-year-period.

Figure one (resp. two) displays the evolution over time of the total daily (resp. weekly) number of news wires. At first sight, the absence of any clear trend indicates that the news time-series is quite stationary. It could seem a bit singular, as we may have expected the frequency of news arrivals to gradually increase over the years. However, it is a rather encouraging and lucky feature of our data, since it eliminates the risk of spurious results due to a possible simultaneous increase over time of the stock volatility. Another interesting fact is that there is some kind of cycle, with 2000 and 2002 as higher news frequency years and 1999, 2001 and 2003 as lower news frequency years. Not surprisingly at all, the Christmas-New Year period experiences a news arrivals rate much below the average. And finally, a close look at Figure one ascertains the presence of weekly seasonality in the data without a shadow of doubt.

Figure three focuses on this weekly seasonality and shows not surprisingly that the average number of news announcements released during the week-end is much lower than the one of the other weekdays.

Fields like company, industry and subject deserve to be examined in detail.

There are 56 different companies in the sample (each of them appears at least once in the CAC40 index over the selected sample period). However, only 40 out of these 56 firms will be analyzed in the subsequent GARCH and MSR econometric models, since some of these companies were created, merged, or acquired during the sample period. Figure 4 shows that news releases intensity was the highest for France Telecom SA (16.6%), Vivendi Universal (14.9%), Renault SA (8.0%), Alcatel SA (7.9%) and Total SA (7.6%)[7]. Such results raise the concern that the two major companies ("major" in terms of "news liquidity", i.e. the number of news releases concerning them) could possibly drive the results and lead us to abusive generalizations. We will deal with this problem by providing individual results for each company and showing that those results are consistent and homogeneous across all firms in our sample.

Concerning industry sectors, they are split into ten general categories and 56 sub-categories. Figure 5 shows that the three sectors with the highest news frequency from 1999 to 2003 were "Financial and Business Services" (48.2%), "Metal, Goods and Engineering" (44.0%) and "Transport and Communication" (38.2%).

Subjects have also been divided into 5 general categories and 105 more specific ones. From Figure 6 one can observe that the 7 most common news subjects among the

[7] Percentages do not sum up to 100%, because one news release often concerns more than one company, industry or subject. It follows that categories are not exclusive (but they are exhaustive).

105 more specific categories are mergers and acquisitions (M&A) (25.5%), earnings projections (19.4%), earnings (10.8%), funding/capital (5.4%), analyst comment/recommendation (5.2%), contracts/orders (4.1%) and regulation/government policy (3.0%). According to Andersen (1996), we expect the impact of various news releases to be different across categories. One of the purposes of the subsequent econometric analysis will be to give some empirical support to such a theoretical prediction.

Table 1 provides an extensive set of summary statistics (mean, min, max, standard deviation, skewness, excess kurtosis and sum), for the whole sample and by company (panel A), by industry (panel B), by subject (panel C) and by timing (panel D). Daily news frequency ranges from 0 to 142 news releases a day, with an average and a standard deviation of respectively 58.50 and 20.12. The skewness and excess kurtosis (moments of order three and four) are close to zero and indicate that the total news probability distribution is asymptotically roughly Gaussian.

Panels A, B, C and D are quite interesting, since they offer valuable information about possible cross-sectional differences in the news arrival rate. One of the main conclusions that can be drawn from the analysis of Table 1 is the following: the greater the number of news releases concerning a company (or an industry, or a subject), the higher its standard deviation and the lower its skewness and kurtosis. The relatively high level of kurtosis experienced by both Carrefour and Promodes is the direct consequence of the very strong and common outlier on August 30th, 2000 (about 70 press releases on that particular day), caused by the announcement of the merger between Carrefour and Promodes.

4. The results

4.1 GARCH framework

The only difference between the basic equation of Kalev and al. (2004)'s regression and ours resides in the introduction of the market return in equation (1), which allows us to control for macroeconomic shocks and to focus on the firm-specific variance of the stock returns. The results reported in Table 2 for each of the forty firms of the sample[8] are largely consistent with Kalev and al. (2004)'s main conclusion: the inclusion of the news variable in the conditional variance equation of the GARCH model reduces volatility persistence ($\alpha + \beta$) for most companies of the sample, even though the magnitude of this reduction is highly variable across firms. The volatility persistence without the inclusion of the news proxy in the equation is close to one for all companies, which is symptomatic of the presence of persistence in the stock volatility, while the same volatility persistence once the news variable is introduced in the regression ranges from a very low 0.1555 for CapGemini (-84%) to a hardly unchanged 0.9999 for Alstom. The magnitude of the reduction in volatility persistence is of course closely related to the significance of the coefficient of the news variable: the higher the significance of the impact of news on volatility, the stronger the decrease of the volatility persistence. Coefficients of the news variable turn out to be positive in all cases and significant for 17 companies out of 40. We will show in a next section of this article that the MSR model performs in fact much better. Note also that the coefficient of the market return is positive and highly significant for all companies, which totally justifies its introduction into the GARCH equations.

4.2 MSR framework: the two-state market model.

For the many reasons brought up in section 2, an MSR framework is expected to be a more appropriate specification and to fit much better than a GARCH model. As it was

[8] Note that we encountered convergence problems with 5 companies (Alcatel, Total, STMicroelectronics, Société Générale and Thalès). This kind of problems is typical of GARCH models.

thoroughly explained in the same section, such an approach leads us to develop a two-state market model that provides us with the (daily) probabilities of being in a high volatility regime on a specific day for each stock of the sample. Once these probabilities have been obtained, it is possible to regress them over our explicative and control variables, company by company. Figure 7 offers first intuitive evidence of the close relation between the probability of being in a high volatility regime and the daily number of news releases. It displays in parallel the time-series evolution of the probability of being in a high volatility state and the daily news frequency for Vivendi Universal, one of the largest firms of the sample. It clearly results from this figure that peaks of volatility generally occur simultaneously with peaks of public information, which is a clue that stock volatility and public information are quite strongly correlated. The next regressions aim at investigating this expected link more formally and precisely.

"Firm by firm" analysis

At this stage, we face two alternatives: either we transform the probability of being in a high volatility regime into a dummy variable and we regress this dummy over our variables of interest through a probit regression, or we directly use the estimated probability of being in a specific volatility state as the dependent variable in a classical WLS regression. However, in that latter case, we have to take into account the inherent heteroscedastic nature of our regression models in order to build correct inferences. We therefore follow the procedure presented in the appendix A (a kind of "weighted least squares regression", hereafter denoted WLS), which leads us to use as a dependent variable the logistic transform of the estimated probability.

Probit regression

In this time-series regression, we transform the state regime into a dummy variable $D_{i,t}$ equal to 1 when the probability of being in the high variance regime at a specific date *t* is greater than 50% and 0 otherwise.

$$D_{i,t} = \beta_{i,0} + \beta_{i,1} N_{i,t} + \beta_{i,2} N_{i,t}^2 + \sum_{j=1}^{J} \delta_{i,j} X_{i,j,t} + \eta_{i,t} \tag{9}$$

where *i* and *t* represent respectively the security i and the time interval *t*;

$D_{i,t}$ = 1 if $P\lfloor S_{i,t} = 2 \rfloor \geq 50\%$ (« high variance regime »)

0 otherwise (« low variance regime ») ;

$N_{i,t}$ stands for the number of news releases specific to security *i* over time interval *t*;

$X_{i,j,t}$ stands for the jth control variable.

For we can a priori suspect the relation between our stock volatility proxy and the daily number of news releases to be non-linear and to exhibit some quadratic pattern, we add the square number of news announcements to the equation. A positive (resp. negative) and significant coefficient of such a variable would reveal the presence of convexity (resp. concavity) in the form of the function relating the news arrival intensity to the stock volatility.

The inclusion of relevant control variables in our regression is necessary in order to test the robustness of the impact of the firm-specific news releases on the conditional variance of the security. The conditional variance of the market portfolio is our main control variable, since all volume proxies (volume levels, differences and log) turned out to be insignificant and thus rather irrelevant on a daily scale. This is totally consistent with Kalev and al. (2004)'s results where volume and first of the week variables are shown to be insignificant on a daily scale. Last but not least, the conditional variance of the sector index is not yet introduced here seeing that some stocks of the sample are the major constituents of their respective sector index, which logically induces an annoying strong correlation between the dependent variable and the control variable. Nevertheless, the sector variable will be well taken into account in the later panel study. Both conditional variances of the market portfolio and of the sector index are estimated via a GARCH (1,1) model.

Table 3 reports the results of the probit regression. The coefficient of the news variable is almost always significantly positive, whereas the coefficient of the square news variable is generally negative and significant, which seems to indicate that the marginal impact of news announcements on the stock volatility is positive but its magnitude decreases as the number of news releases increases. Not surprisingly, the market variance coefficient appears to be positive and significant as the stock volatility is known to convey some market component. Such a control variable allows us to account for all macroeconomic shocks that affect the stock volatility and that are not covered by our set of firm-specific press releases. Finally, note the relatively high cross-sectional variability of the R-square of the regressions. The quality of the R-square is of course mainly driven by the significance of the market variance and news variable coefficients that look like quite heterogeneous and variable across companies.

"Weighted least squares" (WLS) regression for proportion data

Rather than transforming the daily probability of being in the high volatility state into a simple dummy variable and losing some precious and valuable information in the process, it could be more judicious to directly use the above probability as the dependent variable of the time-series regression and to apply the procedure detailed in appendix A (a kind of weighted least squares regression). Note that the only difference with equation (9) of the probit regression is the nature of the dependent variable. The equation of the WLS regression is then as follows:

$$P\left[S_{i,t}=2\right]=\beta_{i,0}+\beta_{i,1}N_{i,t}+\beta_{i,2}N_{i,t}^{2}+\sum_{j=1}^{J}\delta_{i,j}X_{i,j,t}+\eta_{i,t} \tag{10}$$

In such a framework where the dependent variable is a probability rather than a simple dummy, we have to run a weighted logistic regression in order to account for the heteroscedasticity in the variables (see Appendix A). The obtained coefficients are therefore not interpretable as such and they do not measure anymore the marginal impact of a news release, as opposed to those of the probit regression. On the other hand, it is still possible to say something about the significance of the coefficients.

Table 4 displays the results obtained with this second approach. They are very similar to those provided by the probit regression, although they appear to be still more significant and convincing. The coefficient of the news variable (resp. square news variable) is always positive (resp. generally negative) and significant in 38 (resp. 26) cases out of 40. The positive but marginally decreasing impact of public information on stock volatility seems to be demonstrated for most of the companies over the sample period. The other observations made on the basis of the results of the previous probit regression remain valid.

Panel-data analysis

The main advantage of the panel analysis is to work globally and to analyze the combination of time and cross-sectional effects. Modeling in this setting calls for some complex stochastic specification. We therefore use the most common techniques: the fixed effects and random effects approaches.

The fixed effects approach takes the constant model to be a group-specific term in the regression model. The constant is different for any security. It should be noted that the term "fixed" as used here indicates that the constant term does not vary over time.

The random effects approach assumes that the residual is a group specific random element. In this model, the constant is the same for all securities included in the sample.

The crucial distinction between these two cases is whether the unobserved individual effect embodies elements that are correlated with the regressors in the model, not whether this effect is stochastic or not. In our case, the fixed effects model is probably better specified. Indeed, from the relatively high heterogeneity across stocks observed in the individual regressions, we may think that the fixed effects approach fits better, since it assumes the constant to be different for each stock. Moreover, from a statistical perspective, Hausman's test (see the bottom of Table 5) indicates that the fixed effects model performs much better than the random effects one.

Table 5 presents the results for fixed effect and random effects panel approaches. By comparison with the previous firm by firm probit and WLS regressions, the explicative and control variables are unchanged, but the introduction of the conditional variance of the sector index as a new control variable. Both approaches give very similar results: all coefficients are highly significant. Market and especially sector variances seem to strongly affect the stock volatility, and the frequency of news arrivals is shown to have a positive but marginally decreasing impact on the stock volatility. On average, the first news release of the day tends to increase the probability of being in a high volatility regime by roughly 2 basis points, while 50 news releases increase this same probability by about 43 basis points. Naturally, as it could have already been observed in the individual regressions, the same number of press releases will not impact the same way the volatility of small and big firms. One news wire about Vivendi Universal or France Telecom is of course expected to have much less impact on the volatility of these firms than one news release concerning Valeo for example. Finally, it is worth mentioning that our news variable and two control variables explain altogether not less than 26% of the stock volatility in the fixed effects model.

Quite undoubtedly, the level of firm-specific news intensity clearly influences the conditional variance of returns. Such a conclusion brings strong support in favor of the Mixture of Distribution Hypothesis.

Analysis refinement: categorization by topic and by timing

Public information is not homogeneous. We must consider the possibility that a news release about merger and acquisition does not affect the stock volatility the same way as an earnings announcement for example. Similarly, information released during trading hours may not have the same impact on the stock volatility as information released during non-trading hours. From a theoretical perspective (see Andersen (1996)), the topic and the timing of each news release matter. We check this point by running the same regression as above on a limited category of news releases. We define two timing categories: news released during *Trading Hours* and news released during *Non-Trading Hours* (mainly week-end and overnight); and we retain the seven most common subject categories: *Merger and Acquisition (M&A)*, *Earnings*, *Earnings Projections*, *Analyst Comments and Recommendations*[9], *Funding/Capital*, *Regulation and Government Policy*, and, finally, *Contract.* We also choose to add an eighth category named *Earnings Projections clean* which is composed of all *Earnings Projections* except those concerning *Share Price Movement/Disruptions* in order to avoid any endogeneity problem.

Results are synthesized in Table 6 and Figure 8. If the regressions are run individually for each category, with the conditional market and sector variances as the usual control variables, all coefficients are significant at the 5%-level. With a marginal impact of 9.1 points (0.091) on the stock volatility, the *Regulation and Government Policy* category seems to be the critical news category of our sample. *Analyst Comments*

[9] Notice that « analyst comments and recommendations » is actually a subset of « earnings projections ». For obvious reasons of colinearity problems, these two variables will not be introduced together into the multivariate regressions of the next section.

and Recommendations and *Earnings Projections (clean)* come after with coefficients of respectively 6.8 and 4.9. *Contract* (3.4), *M&A* (3.4) and *Earnings* (3.0) follow a bit further and finally comes *Funding and Capital* with a poor 1.2. The high informativeness of the *Regulation and Government Policy* category was not particularly expected. Is it a French specificity due to the important weight of the state in the economy and the society, or does such a type of news impact the same way the stock volatility in other national contexts? This issue could be investigated in further research. Finally, it is interesting to notice that *Earnings Projections* (especially *Analyst Comments and Recommendations*) turn out to be more informative to investors than *Earnings* announcements themselves, since the former induce a greater increase in the stock volatility than the former. Investors and traders appear to value substantially analyst advice and comments.

From a timing perspective, news announcements released during *Trading Hours* seem to have a stronger effect (3.5) on the stock volatility than news announcements released during the week-end or overnight (2.8). This "timing" difference becomes still more impressive if both variables are introduced together in the same regression. First column of Table 7 reports a coefficient of 8.5 for the *Trading Hours* category, weighted by the number of trading hours by day, as opposed to a relatively low coefficient of 1.9 for the *Non-Trading Hours* category, weighted by the number of non trading hours by day. Many reasons can be advanced to explain such a result. It might be that news categories with the strongest impact on volatility (government policy, earnings projections, etc.) occur principally during trading hours; or that investors, consciously or not, don't give the same weight and relevance to information published during trading hours and information published at another time. At this stage, it is too early to definitively answer this question.

We also run a regression with all "category variables" together, except *Analyst Comments and Recommendations* and *Earnings Projections clean* because of obvious problems of colinearity. The last two columns of Table 7 generally confirm the results displayed in Figure 8, though three types of news, *Earnings*, *Funding and Capital* and *Contract* appear now to be insignificant, while *Regulation and Government Policy*, *M&A* and *Earnings Projections* remain highly significant.

Informational content and asymmetry issues

If the daily number of news releases has been proven to be a quite fair proxy of the public information available on the market and to significantly drive the conditional volatility, we still may regret that this variable does not tell anything about the impact of the informational content of news. Indeed, we can legitimately expect that an announcement delivering highly valuable information for investors will induce a stronger increase in the conditional and contemporaneous volatility than any less important piece of news. Our objective is to bring some answer to the following question: "do news releases with different informational contents have different impact on the conditional volatility?" We opt for the absolute value of the daily abnormal stock return as our proxy of the informational and value content of all the public information released each day ("*ARlevel*"), even though we are well aware that this noisy proxy encompasses many more elements and effects than the value effect of information arrivals alone. We also compute a measure of informational content per news release by dividing the variable defined above by the number of announcements of the day ("*ARlevel/NbNews*").

Another issue is brought up by previous empirical research linked to the vast stream of literature concerned with the development of GARCH models. Works such those of Engle and Ng (1993) have concluded to the asymmetry of the volatility response to news. Bad news appears to induce a larger increase in conditional volatility than good one. Accounting for this differentiated effect seems therefore of prime of importance. We aim at answering the following question: "does good and bad news impact the conditional volatility the same way? Or is there any significant asymmetric effect?" Concretely, we simply use the sign of the daily abnormal return, defined as the difference between the market return and the stock return, to classify news into good and bad news[10].

In order to gain some insight of the asymmetry and "informational content" patterns across our two volatility regime, Table 8 provides some basic and nevertheless interesting descriptive statistics. It shows not surprisingly that the average abnormal return in absolute value (*ARlevel*) and the average informational content (*ARlevel/NbNews*) is greater in the high volatility regime than in the low one. More unexpected is the slightly higher percentage of positive AR in the high volatility regime than in the low one, since that existing literature reports that negative innovations usually have a stronger impact on the conditional volatility that the positive ones. For the same reason, the higher average abnormal return of the high volatility regime is quite surprising.

After this first and superficial look at some specificities of our sample, we undertake a more formal analysis and run the following regressions:

$$P\left[S_{i,t}=2\right]=\beta_{i,0}+\beta_{i,1}SignAR_{i,t}+\beta_{i,2}N_{i,t}+\beta_{i,3}N_{i,t}^{2}+\beta_{i,4}SignAR_{i,t}*NbNews$$
$$+\sum_{j=1}^{J}\delta_{i,j}\times X_{i,j,t}+\eta_{i,t} \qquad 11)$$

$$P\left[S_{i,t}=2\right]=\beta_{i,0}+\beta_{i,1}SignAR_{i,t}+\beta_{i,2}ARlevel_{i,t}+\beta_{i,4}SignAR_{i,t}*NbNews$$
$$+\sum_{j=1}^{J}\delta_{i,j}\times X_{i,j,t}+\eta_{i,t} \qquad (12)$$

where $SignAR = 1$ if AR is positive

$ARlevel = |R_i - R_m|$

The four first columns of Table 9 give the results of these regressions. Taken alone, the informational content variable is highly significant, more than *NbNews*. It is

---

[10] Notice that with such a classification rule, all news releases of the same day will be classified in the same category.

already a first clue that informational content matters at least as well as the frequency of news releases. Evidence in favor of possible asymmetry effect is more mixed: when the coefficients of *SignAR* alone or *SignAR* conjugated with the explicative variable (*NbNews* or *ARlevel*) are significant, they are negative and thus in line with the previous literature on the topic.

The daily abnormal return is an interesting proxy of the informational content of all publications released during a day, but it remains a collective measure rather than an individual one, since it does not take into account the information arrivals at all. That is why we choose to also introduce an individual measure of informational content that controls for volume effects, that is, *ARlevel/NbNews*. Regressions 5 and 6 show that with such an explicative variable, asymmetry effects completely vanish, since the coefficients of *SignAR* and *SignAR*InfCont* turn out to be completely insignificant. We understand this new result as such: one good and one bad announcements with identical informational content affect the conditional volatility the same way. Asymmetry between good and bad news disappears once news informational content is accounted for.

5. Conclusion

Without a shadow of doubt, research works aiming at accurately identifying and understanding the link between information diffusion and stock prices are of prime importance and interest, as much for the academic world as they are for business corporations. In this general context, the main objective of our study is to investigate the impact of the news arrivals intensity on the stock price volatility. We use daily prices and Reuters news releases concerning all stocks pertaining to the CAC40 index over the period ranging from January 1, 1999 to December 31, 2003. We first begin with a partial replication of the Kalev and al. (2004)'s GARCH model and we find, like them, that public information explains a large part of the daily volatility persistence. We apply a Markov switching regression (MSR) model to the data in order to determine the daily probability of the volatility regime to be in a high level state or a low level state. Our developed MSR model presents many interesting features: its structure is more general than the GARCH models' one, they more often converge towards a stable solution and they implicitly rest on a theoretical argument: the Mixture of Distribution Hypothesis. We find for each stock a positive and (5%-)significant relation between the daily number of news and the daily probability of being in a high volatility state. The panel analysis confirms these results since the above detected positive relation remains highly significant, even when controlling for market volume and conditional sector and market variances. The daily number of news alone explains not less than 14% of the probability of being in a high variance regime, and 26% in conjunction with the two main control variables.

A positive and significant relation between news arrivals intensity and stock prices volatility seems thus to be reasonable and consistent. Hence it follows that news contains some relevant and valuable information to the market, since news wires have just been proved to significantly impact the stock return distribution. Such a finding is consistent with Roll's idea (and others before him) that the return distribution is mixed ("with news" and "no news" distributions) and driven by news arrivals frequency. It also brings new support in favor of the Mixture of Distribution Hypothesis.

Moreover, further analysis shows that the impact of news intensity on stock volatility is positive but marginally decreasing. The concavity of the relation between number of news releases and stock volatility could indicate a possible phenomenon of saturation, as if too many news didn't bring anymore additional valuable information to the market. By splitting our news sample into different categories according to the timing and the subject, we also find that announcements released during trading hours have a greater impact than news published during non-trading hours. Regarding the content of the news releases, we show that news concerning *Regulation and*

Government Policy has the greatest impact on the stock volatility. *Earnings Projections* (especially the subset of *Analyst Comments and Recommendations*) and *Mergers and Acquisitions* come just after and seem to significantly affect the stock volatility too. The impact of *Earnings* is significant too, but to a limited extent only.

Asymmetry effects reported by the GARCH literature are examined too. Results are globally consistent with previous works, since bad news releases seem to affect more the conditional volatility than good ones, even though their significance is relatively modest. Interestingly, this well-known asymmetric effect vanishes once we account for the informational content of the news releases: good and bad announcements with identical informational content affect the conditional volatility the same way.

Now that the link between news frequency and stock price volatility is quite ascertained, future research could be extended in order to account for private information through the introduction of an appropriate proxy of this phenomenon. From a more dynamic point of view, switching from a daily scale to an intraday scale in order to capture very short-term volatility effects is another potential promising trail of research.

References

Andersen, T. G., 1996, « Return volatility and trading volume: An information flow interpretation of stochastic volatility », *Journal of Finance*, 51, 169-204.

Beaver, William H., 1968, « The information content of annual earnings announcements », *Empirical Research in Accounting*, 6: 67-92.

Berry, T.D., and K. M. Howe, 1993, « Public information arrival », *Journal of Finance*, 49, 1331–1346.

Campbell, J. Y., M. Lettau, B. Malkiel and Y. Xu, 2001, « Have Individual Stocks Become More Volatile? An Empirical Exploration of Idiosyncratic Risk », *Journal of Finance,* 56:1-43, February 2001. Winner of Smith Breeden Prize, 2001.

Chordia, T., R. Roll, and A. Subrahmanyam, 2001, « Evidence on the Speed of Convergence to Market Efficiency », forthcoming: *Journal of Financial Economics*.

Clark, P. K., 1973, « A subordinated stochastic process model with finite variance for speculative prices », *Econometrica*, 41, 135-155.

Easley, D., and M. O'Hara, 2004, « Information and the Cost of Capital », *The Journal of Finance*, 59(4), 1553-1583.

Ederington, L.H., and J. H. Lee, 1993,. How markets process information: News releases and volatility, *Journal of Finance* 48, 1161–1191.

Engle, R., 1982, « Autoregressive conditional heteroskedasticity with estimates of the variance of UK inflation », *Econometrica*, 50, 987-1008.

Engle, R. et V. Ng, 1993, « Measuring and Testing the Impact of News on Volatility », *Journal of Finance*, 48, 1749-1777.

Fama, E. F., 1970, « Efficient Capital Markets : a review of theory and empirical work », *Journal of Finance*, 25, 383-417.

Fama, E., Fisher, L., Jensen, M., Roll, R., 1969, « The adjustment of stock prices to new information », *International Economic Review*, 10, 1-21.

Goyal, A., and P. Santa-Clara, 2003, « Idiosyncratic Risk Matters! », *Journal of Finance*, 58, 975-1007.

Greene, W., 2002, « Econometric Analysis », Fourth edition, *Prentice Hall*.

Hamilton, J., 1989, « A New Approach to the Economic Analysis of Nonstationary Time Series and The Business Cycle », *Econometrica*, 57, 357-84.

Hamilton, J., 1994, « Time series analysis », *Princeton University Press*.

Kalev, P. S., W.-M. Liu, P. K. Pham et E. Jarnecic, 2004, « Public Information Arrival and Volatility of Intraday Stock Returns », *Journal of banking and Finance*, 28(6), 1447-1467.

Krolzig, H.M., 1997, « Markov-Switching Vector Autoregressions », *University of Oxford*, UK Lecture Notes in Economics and Mathematical Systems, Vol. 454, Springer, XIV.

Mitchell, M. L. et J. H. Mulherin, 1994, « The Impact of Public Information on the Stock Market », *Journal of Finance*, 49, 923-950.

Roll, R., 1984, « Orange Juice and weather », *The American Economic Review*, 74(5), pp. 861-880.

Roll, R., 1988, « R-Squared », *Journal of Finance*, 43, 541-566.

Ross, S. A., 1989, « Information and volatility: The no-arbitrage martingale approach to timing and resolution irrelevancy », *Journal of Finance*, 44, 1-17.

Table 1

Summary statistics

NAME	MEAN	MAX	MIN	STD DEV	SKEWNESS	KURTOSIS	SUM
All	**58.50**	**142**	**0**	**20.12**	**0.18**	**0.66**	**76,341**
PANEL A: Daily number of news releases by company							
FRANCE TELECOM	9.72	106	0	8.61	2.89	18.93	12,690
VIVENDI UNIVERSAL	8.71	93	0	8.74	3.06	16.22	11,370
RENAULT	4.69	43	0	4.89	2.27	8.73	6,121
ALCATEL	4.61	38	0	5.13	2.27	7.60	6,017
TOTAL	4.47	56	0	4.30	3.55	26.80	5,835
SOCIETE GENERALE	3.90	74	0	5.18	5.37	49.85	5,089
AVENTIS	3.75	29	0	3.52	2.09	7.00	4,898
BNP PARIBAS	3.64	62	0	4.92	4.71	37.32	4,745
EADS	3.62	61	0	4.38	4.41	38.01	4,720
AXA	2.81	63	0	3.96	5.61	60.22	3,671
SUEZ	2.70	43	0	3.34	3.42	23.19	3,524
STMICROELEC.	2.49	33	0	3.65	3.31	16.62	3,247
LVMH	1.90	32	0	3.27	3.55	17.73	2,484
CANAL PLUS	1.86	40	0	3.63	3.87	22.06	2,422
CARREFOUR	1.83	72	0	3.67	8.98	135.57	2,384
PEUGEOT	1.82	19	0	2.15	2.26	8.66	2,379
CREDIT LYONNAIS	1.81	42	0	2.94	5.38	53.67	2,363
CAP GEMINI	1.78	22	0	2.86	2.72	9.76	2,329
ELF AQUITAINE	1.69	61	0	4.39	5.82	51.73	2,209
DEXIA	1.49	26	0	2.14	3.59	23.00	1,949
ORANGE	1.44	42	0	2.94	4.75	39.88	1,876
PINAULT PRTPS RED.	1.42	29	0	2.81	4.46	27.98	1,853
LAFARGE	1.35	33	0	2.35	5.91	57.75	1,766
BOUYGUES	1.29	31	0	2.57	4.86	36.74	1,680
THALES	1.24	22	0	2.20	4.12	25.62	1,619
LAGARDERE	1.24	36	0	2.66	4.64	34.41	1,618
PARIBAS	1.21	72	0	4.49	8.06	93.55	1,583
CREDIT AGRICOLE	1.20	37	0	2.41	5.29	50.40	1,566
DANONE	1.18	22	0	2.12	3.78	21.99	1,534
ALSTOM	1.10	23	0	2.03	3.91	22.38	1,433
TF1	0.95	18	0	1.80	3.72	20.57	1,236
VEOLIA ENVIRONNEMENT	0.92	27	0	2.41	4.96	32.67	1,199
EQUANT	0.88	18	0	1.98	3.39	14.09	1,146
SCHNEIDER ELECTRIC	0.84	23	0	1.87	4.98	38.27	1,092
SANOFI SYNTHELABO	0.83	18	0	1.80	3.61	17.53	1,087
THOMSON MULTIMEDIA	0.82	22	0	1.98	4.74	31.60	1,065
CASINO	0.78	14	0	1.71	3.81	17.98	1,024
L'OREAL	0.75	26	0	1.67	6.19	62.66	978
ACCOR	0.74	18	0	1.60	4.90	36.38	967
L'AIR LIQUIDE	0.67	32	0	1.91	7.67	85.79	876
MICHELIN	0.65	17	0	1.61	5.21	36.65	844
SAINT GOBAIN	0.61	28	0	1.71	7.06	76.68	800
AGF	0.52	14	0	1.14	3.80	23.36	684
CREDIT COMM.	0.52	29	0	1.81	7.01	71.59	679
VALEO	0.49	13	0	1.31	4.30	23.79	642
PERNOD RICARD	0.45	27	0	1.54	8.12	97.91	586
SODEXHO ALLIANCE	0.44	23	0	1.48	7.71	85.50	575
LEGRAND	0.42	21	0	1.34	7.12	75.33	547
DASSAULT SYSTEMES	0.42	16	0	1.29	5.47	41.11	546
PROMODES	0.41	69	0	2.64	18.55	423.36	537
USINOR	0.29	8	0	0.79	4.49	27.80	383
ARCELOR	0.26	10	0	0.89	4.92	30.44	345
VINCI	0.26	11	0	0.82	5.73	46.94	334
BIC	0.10	7	0	0.50	7.56	74.64	134
SANOFI	0.10	11	0	0.57	10.20	141.22	126
ERIDANIA BEGHIN SAY	0.05	7	0	0.33	10.81	174.04	69

Table 1 (continued)

NAME	MEAN	MAX	MIN	STD DEV	SKEWNESS	KURTOSIS	SUM
PANEL B: Daily number of news releases by industry							
Financial And Business Services	28.18	103	0	12.09	0.63	1.78	36,769
Metal Goods And Engineering	25.71	74	0	12.38	0.46	0.00	33,556
Transport And Communication	22.36	111	0	12.39	1.07	3.39	29,178
Energy	13.14	72	0	9.44	1.51	3.68	17,144
Processing Industries	11.97	74	0	7.49	1.49	5.99	15,615
Metals, Minerals And Chemicals	11.78	50	0	6.90	1.05	1.80	15,379
Distribution, Hotels And Catering	8.16	73	0	5.97	2.38	14.78	10,645
Services And Entertainment	7.73	45	0	5.87	1.78	5.09	10,088
Construction	2.16	33	0	2.89	3.50	21.51	2,815
Agriculture And Forestry	0.35	9	0	0.77	3.39	19.15	463
PANEL C: Daily number of news releases by topic							
M&A	14.94	87	0	8.64	1.56	5.63	19,494
Earnings Projections	11.33	49	0	6.22	0.95	1.79	14,788
Earnings	6.30	41	0	5.60	1.57	3.31	8,223
Funding/Capital	3.14	28	0	2.82	2.36	11.38	4,099
Analyst Comment/Recommendation	3.06	20	0	2.97	1.00	1.18	3,996
Contracts/Orders	2.42	11	0	1.98	1.11	1.55	3,152
Regulation/Government Policy	1.75	15	0	1.96	1.63	3.54	2,278
Share Price Movement/Disruptions	1.38	13	0	1.89	1.98	4.94	1,806
Share Capital	1.29	28	0	1.92	4.54	41.86	1,677
Capacity/Facilities	1.18	8	0	1.31	1.51	2.93	1,541
Plans/Strategy	1.16	16	0	2.13	2.69	8.63	1,508
Marketing	1.13	10	0	1.41	1.79	4.37	1,476
Contracts, Non-government	1.09	9	0	1.41	1.72	3.82	1,419
Corporate Debt Instruments	1.07	13	0	1.44	2.63	12.45	1,393
Monopolies/Antitrust	0.85	8	0	1.25	2.03	5.11	1,113
Management Issues	0.70	8	0	1.01	1.83	4.57	908
Initial Public Offerings	0.67	27	0	1.62	6.63	75.39	877
Management Moves	0.61	8	0	0.95	2.05	5.97	792
Legal/Judicial	0.53	10	0	0.96	2.94	14.07	691
Joint Ventures	0.49	6	0	0.85	2.25	6.32	641
Sales Figures	0.49	5	0	0.95	2.15	4.14	633
Corporate Credit Ratings	0.48	10	0	1.03	3.76	20.14	630
Output/Production	0.47	6	0	0.80	2.18	6.48	608
Dividends	0.44	7	0	0.96	3.07	11.86	570
Privatizations	0.35	8	0	0.81	3.43	16.51	452
New Products/Services	0.35	4	0	0.68	2.33	6.22	452
Labor/Personnel Issues	0.33	11	0	0.73	4.25	39.01	437
Government Contracts	0.33	8	0	0.72	3.46	19.96	428
Stock Listings	0.27	7	0	0.72	3.58	16.42	353
Financing Agreements	0.22	4	0	0.53	2.77	8.96	291
Research/Development	0.21	4	0	0.51	2.77	9.36	280
Defense Contracts	0.21	7	0	0.58	4.18	26.35	269
Divestitures/Asset Sales	0.17	3	0	0.45	2.71	7.44	227
Annual Meetings	0.14	5	0	0.47	4.51	25.50	177
Pricing	0.11	3	0	0.37	3.86	16.93	143
Bankruptcy	0.10	4	0	0.39	5.05	31.83	128
Product Safety	0.09	6	0	0.41	6.48	57.75	122
Market Share	0.09	3	0	0.33	4.23	20.84	115
Licensing Agreements	0.07	2	0	0.28	4.55	21.97	87
External Markets	0.07	3	0	0.26	4.41	23.48	85
Domestic Markets	0.06	2	0	0.25	4.30	19.05	78
Advertising	0.06	2	0	0.25	4.49	21.21	76
Corporate Changes	0.06	2	0	0.25	4.69	23.36	74
Government Aid	0.03	3	0	0.21	7.75	69.83	42
Patents	0.02	2	0	0.16	7.85	68.12	30
Regulatory Bodies	0.02	3	0	0.18	9.73	111.11	29
Conferences/Exhibitions	0.02	3	0	0.17	11.32	156.43	26
Trademarks/Copyrights	0.02	1	0	0.13	7.34	51.96	23
Standards/Standardization	0.01	2	0	0.10	12.31	169.09	12
Earnings Surprises	0.01	1	0	0.10	10.30	104.16	12
Franchises	0.01	2	0	0.10	13.00	189.11	11
Information Technology	0.01	1	0	0.07	13.56	182.14	7
Sales Promotions	0.00	1	0	0.07	14.66	213.33	6
Shareholder-Rights Plans	0.00	1	0	0.06	16.08	256.99	5
Profiles of Companies	0.00	1	0	0.04	25.51	649.99	2
PANEL D: Daily number of news releases - trading hours VS non-trading hours							
Trading hours news	31.99	89	0	12.34	0.21	0.64	41,716
Non-trading hours news	26.55	74	0	10.05	0.53	0.96	34,625

Table 2

GARCH model

We apply the equations (1) and (2) (without news) and (1) and (3) (with news) of the GARCH (1,1) model to our data: 1270 observations for each of the 40 securities. We encountered convergence problems for 5 companies.

			VU	*FT*	*Renault*	*Alcatel*	*Aventis*	*Total*	*STMicro*	*SG*	*AXA*	*Suez*
Garch(1,1)	Without NBNews	**C**	-0.0004	-0.0009	0.0009	-0.0003	0.0002	0.0004	0.0003	0.0009	-0.0002	0.0000
		t-stat	*-0.76*	*-1.38*	*1.54*	*-0.47*	*0.43*	*1.26*	*0.48*	*2.02*	*-0.56*	*-0.04*
		Coeff-MR	1.15	1.58	0.85	1.76	0.84	0.79	1.53	1.13	1.31	0.80
		t-stat	*23.47*	*27.60*	*19.56*	*27.36*	*21.71*	*26.46*	*28.58*	*32.22*	*28.09*	*14.20*
		ω	8.13E-06	1.45E-05	8.00E-06	6.32E-06	1.76E-05	2.95E-07	1.91E-05	7.22E-08	3.17E-06	1.13E-05
		t-stat	*0.69*	*1.49*	*0.93*	*1.04*	*1.07*	*0.54*	*1.45*	*0.09*	*1.20*	*1.64*
		α	0.10	0.08	0.07	0.05	0.09	0.04	0.08	0.03	0.08	0.11
		t-stat	*1.63*	*2.54*	*2.16*	*3.86*	*1.51*	*3.15*	*2.68*	*2.75*	*3.29*	*2.74*
		β	0.89	0.90	0.91	0.94	0.86	0.96	0.89	0.97	0.92	0.86
		t-stat	*11.18*	*21.86*	*19.98*	*62.30*	*8.85*	*70.92*	*20.20*	*80.85*	*33.81*	*16.57*
		α+β	**0.9899**	**0.9840**	**0.9866**	**0.9947**	**0.9491**	**0.9987**	**0.9735**	**0.9999**	**0.9929**	**0.9749**
	With NBNews	**C**	-0.0008	-0.0014	0.0009	-	0.0000	-	-	-	-0.0002	0.0000
		t-stat	*-1.60*	*-2.24*	*1.57*	-	*-0.07*	-	-	-	*-0.48*	*0.02*
		Coeff-MR	1.16	1.49	0.81	-	0.83	-	-	-	1.31	0.80
		t-stat	*30.42*	*26.04*	*20.25*	-	*26.64*	-	-	-	*25.04*	*14.78*
		ω	3.09E-05	8.67E-05	8.77E-05	-	6.68E-05	-	-	-	2.30E-05	7.03E-06
		t-stat	*2.15*	*2.13*	*2.34*	-	*2.20*	-	-	-	*0.50*	*0.86*
		α	0.24	0.20	0.18	-	0.19	-	-	-	0.16	0.13
		t-stat	*4.83*	*3.64*	*3.78*	-	*3.71*	-	-	-	*2.41*	*1.93*
		β	0.20	0.14	0.30	-	0.28	-	-	-	0.64	0.83
		t-stat	*1.89*	*0.97*	*1.97*	-	*1.95*	-	-	-	*1.63*	*8.13*
		α+β	**0.4463**	**0.3416**	**0.4834**	**-**	**0.4721**	**-**	**-**	**-**	**0.8014**	**0.9560**
		λ	2.75E-05	3.87E-05	4.10E-05	-	2.92E-05	-	-	-	1.52E-05	4.19E-06
		t-stat	*4.47*	*3.34*	*3.33*	-	*4.72*	-	-	-	*0.66*	*1.13*

Table 2 (continued)

			LVMH	*BNP*	*Capgem*	*Peugeot*	*Lafarge*	*Danone*	*Schneid.*	*Bouygues*	*PPR*	*Carrefour*
Garch(1,1)	**Without NBNews**	**C**	0.0008	0.0007	-0.0003	0.0007	0.0006	0.0000	0.0006	0.0003	-0.0001	0.0001
		t-stat	*1.74*	*1.80*	*-0.38*	*1.73*	*1.07*	*0.08*	*1.24*	*0.63*	*-0.22*	*0.26*
		Coeff-MR	0.99	0.98	1.40	0.67	0.69	0.44	0.64	0.97	0.90	0.90
		t-stat	*26.30*	*25.71*	*21.79*	*19.76*	*16.30*	*16.78*	*16.18*	*20.63*	*18.51*	*27.03*
		ω	9.69E-06	2.72E-06	2.58E-05	2.03E-05	3.69E-06	1.71E-06	4.27E-06	5.29E-06	9.42E-06	1.66E-06
		t-stat	*1.35*	*0.83*	*1.34*	*1.27*	*1.64*	*1.01*	*0.55*	*1.47*	*1.34*	*0.79*
		α	0.07	0.08	0.04	0.11	0.04	0.06	0.09	0.08	0.04	0.05
		t-stat	*2.44*	*1.70*	*1.91*	*2.17*	*3.90*	*2.32*	*1.66*	*3.88*	*1.83*	*2.22*
		β	0.90	0.91	0.93	0.82	0.95	0.93	0.91	0.92	0.93	0.95
		t-stat	*20.52*	*16.63*	*27.20*	*8.22*	*79.38*	*32.04*	*14.69*	*43.79*	*25.44*	*39.91*
		α+β	**0.9721**	**0.9912**	**0.9724**	**0.9333**	**0.9906**	**0.9944**	**0.9969**	**0.9935**	**0.9752**	**0.9955**
	With NBNews	**C**	0.0004	0.0007	-0.0008	0.0006	0.0007	-0.0002	0.0001	0.0001	-0.0003	-0.0002
		t-stat	*0.85*	*1.77*	*-1.17*	*1.43*	*1.27*	*-0.51*	*0.25*	*0.22*	*-0.69*	*-0.53*
		Coeff-MR	0.98	0.96	1.31	0.67	0.69	0.44	0.62	0.96	0.92	0.87
		t-stat	*31.27*	*25.55*	*24.82*	*22.19*	*16.04*	*16.87*	*15.62*	*22.98*	*26.25*	*29.73*
		ω	2.70E-05	1.15E-05	3.64E-04	3.89E-05	9.09E-07	3.23E-06	1.79E-04	1.24E-04	1.81E-04	9.52E-06
		t-stat	*1.01*	*1.21*	*6.93*	*1.39*	*0.14*	*0.45*	*2.75*	*3.17*	*2.57*	*1.21*
		α	0.09	0.17	0.12	0.17	0.06	0.12	0.21	0.16	0.08	0.15
		t-stat	*1.87*	*3.35*	*2.90*	*4.33*	*1.49*	*0.79*	*4.91*	*4.09*	*1.94*	*3.78*
		β	0.67	0.73	0.03	0.58	0.92	0.84	0.21	0.35	0.14	0.66
		t-stat	*3.30*	*5.51*	*0.60*	*3.11*	*13.35*	*3.68*	*1.17*	*3.10*	*0.72*	*6.68*
		α+β	**0.7632**	**0.9018**	**0.1555**	**0.7427**	**0.9766**	**0.9587**	**0.4206**	**0.5040**	**0.2222**	**0.8026**
		λ	3.00E-05	3.48E-06	2.16E-04	1.97E-05	6.74E-06	8.03E-06	1.24E-04	1.23E-04	5.74E-05	3.11E-05
		t-stat	*1.67*	*0.95*	*8.16*	*1.78*	*0.95*	*0.45*	*3.44*	*4.06*	*4.66*	*3.15*

Table 2 (continued)

		Alstom	*Dassault*	*Sanofi*	*Thales*	*Accor*	*TF1*	*Lagardère*	*Air liq.*	*Michelin*	*L'oréal*
Without NBNews	**C**	0.0001	0.0007	0.0004	0.0002	0.0004	0.0004	0.0006	0.0004	0.0002	0.0000
	t-stat	*0.07*	*0.83*	*0.87*	*0.31*	*0.78*	*0.64*	*1.19*	*0.95*	*0.46*	*-0.05*
	Coeff-MR	0.68	1.26	0.81	0.65	0.81	1.08	0.95	0.77	0.58	0.87
	t-stat	*6.65*	*20.77*	*24.12*	*14.99*	*21.29*	*17.41*	*19.65*	*22.38*	*14.10*	*24.65*
	ω	1.60E-05	1.83E-05	1.44E-06	4.60E-06	3.28E-05	6.26E-06	2.18E-06	1.29E-06	6.39E-05	6.22E-07
	t-stat	*1.16*	*0.38*	*0.44*	*1.42*	*1.58*	*1.03*	*1.18*	*0.95*	*2.82*	*0.34*
	α	0.11	0.06	0.04	0.08	0.11	0.08	0.06	0.04	0.16	0.06
	t-stat	*2.58*	*0.76*	*1.21*	*3.57*	*2.34*	*2.31*	*4.69*	*3.16*	*2.91*	*1.77*
	β	0.90	0.92	0.95	0.92	0.81	0.92	0.93	0.95	0.68	0.94
	t-stat	*29.33*	*7.54*	*22.52*	*43.74*	*8.87*	*27.74*	*69.65*	*62.15*	*8.51*	*26.63*
	α+β	**0.9999**	**0.9821**	**0.9967**	**0.9947**	**0.9206**	**0.9963**	**0.9976**	**0.9952**	**0.8415**	**0.9995**
With NBNews	**C**	-0.0001	0.0001	0.0003	-	0.0003	0.0003	0.0006	0.0004	0.0002	-0.0003
	t-stat	*-0.06*	*0.18*	*0.61*	-	*0.57*	*0.47*	*1.24*	*0.87*	*0.43*	*-0.76*
	Coeff-MR	0.66	1.23	0.81	-	0.81	1.06	0.93	0.76	0.61	0.86
	t-stat	*7.66*	*18.26*	*25.32*	-	*21.85*	*23.20*	*21.36*	*22.16*	*15.91*	*26.42*
	ω	3.63E-07	2.17E-04	3.99E-06	-	9.43E-05	1.24E-06	1.96E-06	1.80E-06	6.99E-05	2.62E-07
	t-stat	*0.02*	*0.83*	*0.65*	-	*2.13*	*0.15*	*0.46*	*0.79*	*2.90*	*0.09*
	α	0.11	0.21	0.09	-	0.18	0.11	0.06	0.04	0.15	0.11
	t-stat	*2.54*	*2.49*	*1.67*	-	*3.99*	*2.21*	*2.88*	*2.58*	*3.80*	*3.02*
	β	0.89	0.50	0.87	-	0.49	0.86	0.92	0.94	0.55	0.86
	t-stat	*20.13*	*1.24*	*10.24*	-	*2.94*	*11.30*	*24.42*	*31.90*	*6.13*	*16.59*
	α+β	**0.9999**	**0.7171**	**0.9602**	**-**	**0.6727**	**0.9655**	**0.9803**	**0.9848**	**0.7030**	**0.9713**
	λ	2.19E-05	1.91E-04	1.45E-05	-	5.44E-05	3.23E-05	7.55E-06	3.62E-06	6.54E-05	1.26E-05
	t-stat	*1.05*	*0.88*	*1.32*	-	*2.07*	*1.15*	*1.26*	*1.00*	*3.11*	*2.10*

Table 2 (continued)

			Valeo	Sodexho	Casino	Pernod	St-Gobain	Equant	AGF	Arcelor	Vinci	Bic
Garch(1,1)	Without NBNews	C	-0.0001	0.0000	0.0002	0.0007	0.0010	0.0004	0.0000	0.0008	0.0006	0.0004
		t-stat	*-0.23*	*-0.06*	*0.46*	*1.34*	*2.19*	*0.43*	*-0.12*	*1.18*	*1.12*	*0.65*
		Coeff-MR	0.62	0.59	0.48	0.23	0.79	1.31	0.55	0.62	0.28	0.20
		t-stat	*14.34*	*7.76*	*12.47*	*4.71*	*10.88*	*18.36*	*11.30*	*11.98*	*9.79*	*4.85*
		ω	2.62E-05	2.97E-06	2.01E-06	2.33E-05	1.23E-05	2.89E-05	2.40E-06	2.09E-05	1.98E-07	2.59E-05
		t-stat	*1.03*	*0.74*	*1.10*	*1.05*	*1.37*	*2.06*	*1.33*	*1.87*	*0.22*	*1.32*
		α	0.10	0.04	0.03	0.14	0.19	0.11	0.06	0.06	0.01	0.11
		t-stat	*1.90*	*3.40*	*3.46*	*1.74*	*1.84*	*3.41*	*3.66*	*3.54*	*2.00*	*2.08*
		β	0.85	0.96	0.97	0.82	0.81	0.87	0.93	0.91	0.99	0.84
		t-stat	*8.96*	*124.38*	*86.98*	*7.18*	*9.51*	*26.38*	*46.67*	*32.90*	*121.78*	*10.21*
		α+β	**0.9522**	**0.9970**	**0.9927**	**0.9580**	**0.9999**	**0.9797**	**0.9927**	**0.9657**	**0.9988**	**0.9506**
	With NBNews	C	-0.0003	0.0004	-0.0001	0.0006	0.0008	-0.0001	-0.0001	0.0008	0.0005	0.0003
		t-stat	*-0.62*	*0.75*	*-0.24*	*1.14*	*1.58*	*-0.16*	*-0.33*	*1.18*	*0.63*	*0.50*
		Coeff-MR	0.62	0.65	0.51	0.22	0.78	1.30	0.55	0.62	0.30	0.18
		t-stat	*15.71*	*13.59*	*17.37*	*5.05*	*19.99*	*18.83*	*11.26*	*11.98*	*9.50*	*4.49*
		ω	5.44E-05	1.57E-04	1.39E-04	2.82E-06	9.19E-05	3.88E-05	3.24E-07	2.09E-05	9.89E-05	1.16E-04
		t-stat	*2.01*	*4.52*	*3.21*	*0.22*	*1.38*	*1.60*	*0.18*	*1.87*	*0.86*	*1.06*
		α	0.16	0.21	0.11	0.12	0.18	0.15	0.06	0.06	0.21	0.21
		t-stat	*4.84*	*4.54*	*2.66*	*1.33*	*4.07*	*3.05*	*3.22*	*3.54*	*2.68*	*2.06*
		β	0.66	0.30	0.18	0.84	0.45	0.80	0.93	0.91	0.48	0.51
		t-stat	*6.96*	*3.77*	*1.10*	*6.52*	*1.83*	*10.33*	*36.90*	*32.90*	*1.08*	*1.53*
		α+β	**0.8214**	**0.5153**	**0.2889**	**0.9585**	**0.6269**	**0.9431**	**0.9941**	**0.9657**	**0.6894**	**0.7190**
		λ	8.65E-05	2.26E-04	8.45E-05	1.01E-05	1.26E-04	4.28E-05	3.68E-06	0.00E+00	6.61E-05	2.68E-04
		t-stat	*2.72*	*4.05*	*4.45*	*1.43*	*1.66*	*1.23*	*1.16*	*0.00*	*0.62*	*1.31*

Table 3

Probit regressions

We run a probit regression with 1270 observations for each of the 40 securities. The dependent variable is $D_{i,t} = 1$ if $P\left[S_{i,t} = 2\right] \geq 50\%$ (« high variance regime »), 0 otherwise (« low variance regime »). C stands for the constant. NbNews is the daily number of news releases concerning a given company. NbNews² is the square NbNews in order to account for a possible quadratic effect. MarketVar is the conditional variance of the market portfolio estimated via a GARCH (1,1). See equation (9).

		VU	*FT*	*Renault*	*Alcatel*	*Aventis*	*Total*	*STMicro*	*SG*	*AXA*	*Suez*
Probit (Y=dummy)	**C**	-2.89	-2.64	-0.21	-2.86	-0.96	-0.28	-0.05	-0.11	-2.36	-3.07
	t-stat	*-21.11*	*-17.30*	*-2.64*	*-18.43*	*-10.89*	*-3.06*	*-0.83*	*-1.34*	*-17.79*	*-19.51*
	Coeff-NBNEWS	0.0946	0.0889	0.1032	0.1121	0.1380	0.0983	0.0502	-0.0097	0.0887	0.0106
	t-stat	*7.65*	*5.97*	*6.75*	*4.85*	*6.02*	*3.65*	*2.65*	*-0.35*	*5.01*	*0.30*
	Coeff-NBNEWS²	-0.0007	-0.0011	-0.0025	-0.0014	-0.0042	-0.0003	-0.0018	0.0054	-0.0012	0.0032
	t-stat	*-2.90*	*-3.17*	*-3.91*	*-1.72*	*-3.17*	*-0.15*	*-1.87*	*2.56*	*-2.77*	*1.57*
	Coeff-MarketVar	3048.93	2121.78	536.28	2546.57	1348.79	257.61	-47.35	469.33	8200.15	6155.37
	t-stat	*13.25*	*10.63*	*3.01*	*11.54*	*7.43*	*1.46*	*-0.27*	*2.64*	*15.61*	*16.48*
	R²	0.44	0.16	0.05	0.24	0.09	0.06	0.01	0.04	0.38	0.58

		LVMH	*BNP*	*Capgem*	*Peugeot*	*Lafarge*	*Danone*	*Schneid.*	*Bouygues*	*PPR*	*Carrefour*
Probit (Y=dummy)	**C**	-1.10	-1.00	-1.51	-0.98	-0.24	-0.18	-0.94	-0.76	-1.56	-0.69
	t-stat	*-14.96*	*-11.19*	*-18.81*	*-12.87*	*-3.63*	*-2.71*	*-13.92*	*-11.20*	*-20.06*	*-10.23*
	Coeff-NBNEWS	0.1772	0.1241	0.1818	0.0861	0.1034	0.1580	0.2000	0.2782	0.0828	0.1525
	t-stat	*7.85*	*5.24*	*6.31*	*2.40*	*3.58*	*5.01*	*5.78*	*10.40*	*3.43*	*8.44*
	Coeff-NBNEWS²	-0.0053	-0.0003	-0.0072	0.0014	-0.0014	-0.0058	-0.0083	-0.0091	-0.0013	-0.0018
	t-stat	*-4.35*	*-0.22*	*-3.30*	*0.35*	*-0.73*	*-2.22*	*-3.23*	*-6.59*	*-1.04*	*-4.12*
	Coeff-MarketVar	1443.70	1655.81	2698.03	1287.57	866.40	-516.41	838.55	750.30	2351.80	244.55
	t-stat	*8.05*	*8.37*	*12.50*	*7.18*	*4.84*	*-2.88*	*4.69*	*4.29*	*12.32*	*1.38*
	R²	0.11	0.17	0.20	0.07	0.04	0.04	0.05	0.12	0.17	0.08

Table 3 (continued)

		Alstom	Dassault	Sanofi	Thales	Accor	TF1	Lagardère	Air liq.	Michelin	L'oréal
Probit (Y=dummy)	**C**	-2.19	-0.85	-0.05	-0.17	-1.44	-0.36	-0.77	-1.06	-1.92	-0.20
	t-stat	*-22.07*	*-13.32*	*-0.79*	*-2.51*	*-19.12*	*-5.83*	*-11.41*	*-15.13*	*-21.23*	*-3.10*
	Coeff-NBNEWS	0.1555	0.0410	0.1675	0.0662	0.1874	0.1431	0.2146	0.1951	0.2179	0.1552
	t-stat	*3.67*	*0.75*	*4.11*	*2.86*	*4.26*	*3.83*	*8.84*	*5.58*	*4.19*	*3.97*
	Coeff-NBNEWS²	-0.0042	0.0005	-0.0108	-0.0016	-0.0087	-0.0077	-0.0056	-0.0055	-0.0049	-0.0055
	t-stat	*-1.25*	*0.09*	*-2.52*	*-1.09*	*-2.24*	*-2.12*	*-4.35*	*-3.15*	*-1.10*	*-2.00*
	Coeff-MarketVar	2475.55	1345.08	918.05	40.51	1735.52	387.29	99.65	715.30	1871.20	939.36
	t-stat	*12.37*	*7.52*	*4.75*	*0.24*	*9.32*	*2.23*	*0.54*	*3.94*	*9.56*	*5.01*
	R²	0.20	0.05	0.03	0.01	0.10	0.02	0.09	0.04	0.12	0.04

		Valeo	Sodexho	Casino	Pernod	St-Gobain	Equant	AGF	Arcelor	Vinci	Bic
Probit (Y=dummy)	**C**	-0.44	-1.09	-0.84	-0.59	-1.26	-1.09	-1.35	-0.45	-0.24	-0.95
	t-stat	*-6.96*	*-16.32*	*-12.53*	*-9.02*	*-17.72*	*-15.28*	*-18.93*	*-6.67*	*-3.97*	*-14.74*
	Coeff-NBNEWS	0.3476	0.2341	0.2204	0.0478	0.2560	0.1974	0.1755	0.1929	0.0127	0.5735
	t-stat	*5.86*	*4.43*	*4.62*	*2.15*	*6.49*	*4.68*	*2.87*	*2.88*	*0.16*	*2.52*
	Coeff-NBNEWS²	-0.0283	-0.0064	-0.0114	-0.0016	-0.0091	-0.0074	-0.0068	-0.0163	0.0009	0.0058
	t-stat	*-3.73*	*-1.21*	*-2.16*	*-1.37*	*-3.76*	*-1.63*	*-0.67*	*-1.32*	*0.07*	*0.06*
	Coeff-MarketVar	991.41	2048.54	562.59	1063.78	1702.00	2212.50	2670.75	1774.83	-173.00	1355.79
	t-stat	*5.58*	*10.75*	*3.18*	*5.92*	*9.07*	*11.63*	*13.58*	*8.83*	*-0.98*	*7.64*
	R²	0.05	0.14	0.62	0.03	0.11	0.14	0.20	0.08	0.00	0.08

Table 4

WLS regressions

We run a WLS ("weighted least squares", please see appendix A for more details) regression with 1270 observations for each of the 40 securities. The dependent variable is the daily probability of being in the high-variance regime. C stands for the constant. NbNews is the daily number of news releases concerning a given company. NbNews² is the square NbNews in order to account for a possible quadratic effect. MarketVar is the conditional variance of the market portfolio estimated via a GARCH (1,1). See equation (10).

		VU	*FT*	*Renault*	*Alcatel*	*Aventis*	*Total*	*STMicro*	*SG*	*AXA*	*Suez*
GLS (Y=proba)	**C**	-6.63	-4.46	-0.72	-4.59	-1.96	-1.55	-0.69	-0.79	-3.80	-6.96
	t-stat	*-67.26*	*-50.33*	*-2.73*	*-59.58*	*-12.68*	*-4.57*	*-3.32*	*-4.41*	*-16.65*	*-86.27*
	Coeff-NBNEWS	0.1131	0.1090	0.2630	0.0880	0.3292	0.5241	0.2374	0.2313	0.3309	0.0304
	t-stat	*5.70*	*11.23*	*6.10*	*4.49*	*7.07*	*14.92*	*4.30*	*9.77*	*17.23*	*1.08*
	Coeff-NBNEWS²	0.0005	-0.0011	-0.0049	0.0002	-0.0110	-0.0095	-0.0097	-0.0031	-0.0094	0.0043
	t-stat	*1.32*	*-6.09*	*-2.65*	*0.14*	*-3.95*	*-11.04*	*-4.76*	*-6.32*	*-10.32*	*3.13*
	Coeff-MarketVar	4558.33	4635.06	2458.20	5265.58	2145.86	-191.99	1827.45	1633.88	9067.96	10962.70
	t-stat	*9.19*	*14.68*	*4.95*	*16.63*	*6.47*	*-0.28*	*3.50*	*4.07*	*42.61*	*74.68*
	R²	0.38	0.25	0.08	0.24	0.11	0.22	0.03	0.11	0.65	0.83

		LVMH	*BNP*	*Capgem*	*Peugeot*	*Lafarge*	*Danone*	*Schneid.*	*Bouygues*	*PPR*	*Carrefour*
GLS (Y=proba)	**C**	-2.93	-2.70	-3.42	-1.98	-0.80	-1.19	-2.08	-2.49	-2.55	-2.04
	t-stat	*-25.04*	*-13.45*	*-36.57*	*-15.83*	*-3.70*	*-6.24*	*-18.29*	*-12.79*	*-30.31*	*-10.40*
	Coeff-NBNEWS	0.5750	0.2184	0.3714	0.2257	0.4441	0.6424	0.5089	0.7968	0.1629	0.5230
	t-stat	*12.46*	*23.40*	*6.74*	*3.74*	*16.40*	*6.30*	*7.15*	*21.05*	*4.60*	*14.83*
	Coeff-NBNEWS²	-0.0159	-0.0015	-0.0147	-0.0021	-0.0046	-0.0318	-0.0242	-0.0275	-0.0031	-0.0094
	t-stat	*-5.96*	*-14.59*	*-3.14*	*-0.39*	*-5.82*	*-3.24*	*-4.19*	*-20.63*	*-1.45*	*-6.72*
	Coeff-MarketVar	4673.18	5125.65	6007.60	3070.40	2406.27	-533.23	2322.94	2429.82	3990.37	-759.61
	t-stat	*13.88*	*17.08*	*26.27*	*9.19*	*5.94*	*-1.12*	*6.30*	*4.27*	*16.72*	*-1.38*
	R²	0.38	0.55	0.54	0.10	0.72	0.05	0.08	0.29	0.23	0.46

Table 4 (continued)

GLS (Y=proba)	Alstom	Dassault	Sanofi	Thales	Accor	TF1	Lagardère	Air liq.	Michelin	L'oréal
C	-5.84	-2.64	0.91	-0.76	-3.52	-1.26	-3.85	-2.26	-4.25	-0.28
t-stat	*-46.55*	*-21.45*	*2.49*	*-3.28*	*-30.89*	*-5.29*	*-16.06*	*-18.19*	*-37.64*	*-1.20*
Coeff-NBNEWS	0.0431	0.3135	0.5956	0.3277	0.3037	0.6804	0.4044	0.4555	0.5403	0.1480
t-stat	*0.53*	*2.53*	*5.38*	*4.13*	*3.25*	*4.69*	*6.80*	*7.34*	*4.75*	*3.30*
Coeff-NBNEWS²	0.0379	-0.0020	-0.0211	-0.0082	-0.0067	-0.0255	0.0077	-0.0116	-0.0312	0.0090
t-stat	*5.53*	*-0.18*	*-2.53*	*-1.60*	*-0.75*	*-1.85*	*2.75*	*-3.51*	*-2.81*	*5.39*
Coeff-MarketVar	7675.52	5031.74	1305.53	794.34	5213.88	-403.04	4683.57	1551.48	4425.09	3412.38
t-stat	*12.48*	*15.28*	*1.82*	*1.42*	*12.50*	*-0.61*	*9.01*	*3.86*	*9.23*	*8.33*
R²	0.23	0.24	0.06	0.04	0.14	0.05	0.58	0.11	0.10	0.80

GLS (Y=proba)	Valeo	Sodexho	Casino	Pernod	St-Gobain	Equant	AGF	Arcelor	Vinci	Bic
C	-1.39	-0.84	-1.65	-1.48	-4.27	-4.78	-3.76	-0.26	0.10	-2.24
t-stat	*-8.13*	*-12.53*	*-12.47*	*-8.68*	*-30.47*	*-21.21*	*-10.69*	*-1.92*	*0.70*	*-11.67*
Coeff-NBNEWS	1.2734	0.2204	0.7200	0.1846	0.5943	4.0426	2.5080	0.3691	0.5528	2.8986
t-stat	*8.91*	*4.62*	*7.88*	*3.05*	*3.86*	*27.50*	*19.29*	*2.92*	*3.24*	*20.72*
Coeff-NBNEWS²	-0.0928	-0.0114	-0.0338	-0.0071	0.0201	-0.2973	-0.2054	-0.0208	-0.0657	-0.2445
t-stat	*-4.05*	*-2.16*	*-4.18*	*-2.36*	*1.37*	*-16.97*	*-12.43*	*-0.92*	*-2.83*	*-11.53*
Coeff-MarketVar	2667.88	562.59	1409.16	3346.12	7854.07	8192.35	4689.97	2812.45	-854.61	4791.39
t-stat	*6.99*	*3.18*	*3.57*	*8.07*	*16.51*	*26.39*	*11.71*	*11.84*	*-2.14*	*9.50*
R²	0.18	0.04	0.14	0.07	0.57	0.67	0.49	0.11	0.01	0.61

Table 5

Panel data analysis

We opt for the fixed effects and random effects approaches. The dependent variable is alternatively the volatility regime dummy and the probability of being in the high volatility regime. C stands for the constant. NbNews is the daily number of news releases concerning a given company. $NbNews^2$ is the square NbNews in order to account for a possible quadratic effect. MarketVar and SectorVar are respectively the conditional variance of the market portfolio and the conditional variance of the sector index. Both of them are estimated via a GARCH (1,1). Hausman's test is largely significant and provides evidence in favor of the fixed effects approach. The panel is composed of 1270 observations for each of the 40 securities.

		Probit (Y=dummy)				*GLS (Y=proba)*			
		1	**2**	**3**	**4**	**1**	**2**	**3**	**4**
Fixed Effects	**C**	-	-	-	-	-	-	-	-
	t-stat	-	-	-	-	-	-	-	-
	Coeff-NBNEWS	0.020291	0.029837	0.029082	0.022915	0.019403	0.028254	0.027561	0.021872
	t-stat	*36.20*	*35.82*	*35.60*	*28.86*	*42.69*	*41.85*	*41.87*	*34.59*
	Coeff-NBNEWS²	-	-0.000283	-0.000279	-0.000286	-	-0.000262	-0.000258	-0.000265
	t-stat	-	*-15.46*	*-15.56*	*-16.55*	-	*-17.69*	*-17.91*	*-19.27*
	Coeff-MarketVar	-	-	432.799	113.1	-	-	396.695	101.74
	t-stat	-	-	*45.28*	*10.70*	-	-	*51.50*	*12.09*
	Coeff-SectorVar	-	-	-	240.37	-	-	-	221.77
	t-stat	-	-	-	*61.87*	-	-	-	*71.68*
	R²	0.1118	0.1159	0.1502	0.2098	0.1375	0.1428	0.1854	0.2603

		Probit (Y=dummy)				*GLS (Y=proba)*			
		1	**2**	**3**	**4**	**1**	**2**	**3**	**4**
Random Effects	**C**	0.319681	0.305827	0.189661	0.144981	0.340982	0.328124	0.221651	0.180393
	t-stat	*13.29*	*12.54*	*7.76*	*5.09*	*42.60*	*14.77*	*9.97*	*6.94*
	Coeff-NBNEWS	0.020217	0.02971	0.028961	0.022828	0.019345	0.028155	0.027468	0.021807
	t-stat	*36.12*	*35.72*	*35.50*	*28.78*	*42.60*	*41.75*	*41.78*	*34.51*
	Coeff-NBNEWS²	-	-0.000281	-0.000278	-0.000285	-	-0.000261	-0.000258	-0.000264
	t-stat	-	*-15.40*	*-15.50*	*-16.49*	-	*-17.62*	*-17.85*	*-19.21*
	Coeff-MarketVar	-	-	432.837	114.189	-	-	396.724	102.505
	t-stat	-	-	*45.28*	*10.81*	-	-	*51.50*	*12.18*
	Coeff-SectorVar	-	-	-	239.58	-	-	-	221.21
	t-stat	-	-	-	*61.73*	-	-	-	*71.56*
	R²	0.0119	0.0133	0.0451	0.0641	0.0154	0.0170	0.0559	0.0803

Hausman's test:	6.8931***	10.815***	10.521***	26.85***	8.1273***	12.014***	11.704***	27.971***

Table 6

Panel data analysis for various types of news: categories are not mixed

We opt for the fixed effects approach. The dependent variable is the probability of being in the high volatility regime. C stands for the constant. NbNews is the daily number of news releases concerning a given company. NbNews² is the square NbNews in order to account for a possible quadratic effect. MarketVar and SectorVar are respectively the conditional variance of the market portfolio and the conditional variance of the sector index. Both of them are estimated via a GARCH (1,1). A first distinction is made between news released during non-trading hours and news released during trading hours. A second distinction is made according to the subject: Mergers and Acquisitions, Earnings, Analyst Comment/Recommendation, Earnings Projections, Funding and Capital, Regulation and Government Policy and, finally, Contracts/Orders. The panel is composed of 1270 observations for each of the 40 securities.

		Dependent Variable : Smooth Probability				
	News category	***All***	***Trading Hours***	***No Trading Hours***	***M&A***	***Earnings***
Fixed Effects	**C**	-	-	-	-	-
	t-stat	-	-	-	-	-
	Coeff-NBNEWS	0.021872	0.034547	0.028395	0.033537	0.030304
	t-stat	*34.59*	*34.13*	*24.06*	*21.53*	*7.91*
	Coeff-NBNEWS(All)²	-0.000265	-0.000640	-0.000685	-0.000464	-0.002623
	t-stat	*-19.27*	*-16.45*	*-13.39*	*-8.21*	*-3.97*
	Coeff-MarketVar	101.74	100.1	94.9941	98.9408	83.2537
	t-stat	*12.09*	*11.91*	*11.22*	*11.68*	*9.80*
	Coeff-SectorVar	221.765	222.516	229.369	235.252	238.493
	t-stat	*71.68*	*72.26*	*73.86*	*76.86*	*77.57*
	R^2	0.2603	0.2603	0.2507	0.2504	0.2438

	News category	***Analysts Comments***	***Earnings Projections (clean)***	***Funding/Capital***	***Reg/Govn***	***Contracts/Orders***
Fixed Effects	**C**	-	-	-	-	-
	t-stat	-	-	-	-	-
	Coeff-NBNEWS	0.067531	0.049423	0.012476	0.09079	0.033923
	t-stat	*9.05*	*17.86*	*2.37*	*15.82*	*3.20*
	Coeff-NBNEWS(All)²	-0.005736	-0.001938	-0.000271	-0.003832	-0.006722
	t-stat	*-2.21*	*-5.34*	*-0.44*	*-4.53*	*-1.58*
	Coeff-MarketVar	202.813	86.9549	85.7048	94.2763	85.1892
	t-stat	*23.41*	*10.27*	*10.08*	*11.10*	*10.02*
	Coeff-SectorVar	209.591	234.757	238.514	238.122	239.342
	t-stat	*65.25*	*76.47*	*77.32*	*77.67*	*77.87*
	R^2	0.2484	0.2479	0.2418	0.2459	0.2418

Table 7

Panel data analysis for various types of news: categories are mixed together

Fixed effects approach (the dependent variable is the daily probability of being in a high volatility regime). The conditional variance of the market portfolio and the conditional variance of the sector index are estimated via a GARCH (1,1). The panel is composed of 1270 observations for each of the 40 securities. The first regression introduces the two timing categories; the second one introduces six subject categories; and the third one introduces the square variables in addition of the variables of the second regression.

	GLS (Y=proba)		
	1	*2*	*3*
Coeff-NBNEWS(Trading Hours)-weighted	0.08483	-	-
t-stat	*26.55*	-	-
Coeff-NBNEWS(Trading Hours)²	-0.00392382	-	-
t-stat	*-11.24*	-	-
Coeff-NBNEWS(Non-Trading Hours)-weighted	0.019167	-	-
t-stat	*9.45*	-	-
Coeff-NBNEWS(Non-Trading Hours)²	-0.00115784	-	-
t-stat	*-8.47*	-	-
Coeff-NBNEWS(M&A)-C181	-	0.016258	0.022737
t-stat	-	*13.83*	*13.36*
Coeff-NBNEWS(M&A)²	-	-	-0.000214
t-stat	-	-	*-3.07*
Coeff-NBNEWS(Earnings)-C151	-	-0.000969117	0.005554
t-stat	-	*-0.36*	*1.36*
Coeff-NBNEWS(Earnings)²	-	-	-0.001369
t-stat	-	-	*-1.97*
Coeff-NBNEWS(Earnings Proj)-C152	-	0.02267	0.038542
t-stat	-	*10.72*	*13.27*
Coeff-NBNEWS(Earnings Proj)²	-	-	-0.002598
t-stat	-	-	*-7.06*
Coeff-NBNEWS(Funding/Kal)-C17	-	-0.0042867	-0.003993
t-stat	-	*-1.16*	*-1.07*
Coeff-NBNEWS(Funding/Kal)²	-	-	-
t-stat	-	-	-
Coeff-NBNEWS(Reg/Govn)-C13	-	0.047598	0.064975
t-stat	-	*9.80*	*10.87*
Coeff-NBNEWS((Reg/Govn))²	-	-	-0.004416
t-stat	-	-	*-4.74*
Coeff-NBNEWS(Contracts)-C33	-	0.010044	0.02001
t-stat	-	*1.66*	*1.90*
Coeff-NBNEWS(Contracts)²	-	-	-0.006001
t-stat	-	-	*-1.42*
Coeff-MarketVar	100.602	99.3205	100.212
t-stat	*11.96*	*11.73*	*11.84*
Coeff-SectorVar	222.005	231.768	231.833
t-stat	*71.76*	*75.20*	*75.23*
R^2	0.2617	0.2530	0.2555

Table 8

Summary Statistics

This table presents some summary statistics for the decomposition regime of the 40 securities studied. The high volatility regime consists of all days where $P\lfloor S_{i,t} = 2 \rfloor \geq 50\%$, while the low volatility regime consists of all remaining days. AR = R_{it} - R_{mt} and ARlevel = $|R_i - R_m|$. InfCont is defined as ARlevel/NbNews and is used as a proxy for the average informational content of all news releases of a given day.

	# Days	Mean-AR	Mean-Arlevel	% AR > 0	# News	Mean - InfCont
High Volatility Regime	18642	0.001001	0.024907	50.45%	48985	0.006940
Low Volatility Regime	32158	-0.000195	0.012403	48.94%	58164	0.003252
Total	50800	0.000244	0.016992	49.50%	107149	0.004605

Table 9

Informational value and asymmetry issues

This table presents various regressions of the smooth probability of being in the high volatility regime on some expected proxies of information content. The purpose of DsignAR is to take possible asymmetry effects into account. It is equal to 1 if AR>0 ("good news") and 0 otherwise "bad news"). NbNews is the daily number of news releases. ARlevel is equal to $|R_i - R_m|$. InfCont is defined as ARlevel / NbNews and is used as a proxy for the average informational content of all news releases of a given day.

		Dependent Variable : Proba smooth					
		1	***2***	***3***	***4***	***5***	***6***
Fixed Effect	**C**	-	-	-	-	-	-
	t-stat	-	-	-	-	-	-
	Coeff-SignAR	-0.00474	-0.001878	0.000919	0.00426	0.001079	0.00011
	t-stat	*-1.72*	*-0.48*	*0.31*	*1.26*	*0.36*	*0.03*
	Coeff-NBNEWS	-	-	0.021711	0.022609	-	-
	t-stat	-	-	*34.48*	*30.05*	-	-
	Coeff-NBNEWS²	-	-	-0.000263	-0.000267	-	-
	t-stat	-	-	*-19.17*	*-19.28*	-	-
	Coeff-ARlevel	9.03672	9.12614	-	-	-	-
	t-stat	*106.48*	*75.64*	-	-	-	-
	Coeff-DsignAR*NBNews	-	-	-	-0.00159	-	-
	t-stat	-	-	-	*-2.18*	-	-
	Coeff-DsignAR*ARlevel	-	-0.168426	-	-	-	-
	t-stat	-	*-1.04*	-	-	-	-
	Coeff-InfCont	-	-	-	-	8.88635	8.77731
	t-stat	-	-	-	-	*53.87*	*37.22*
	Coeff-DsignAR*InfCont	-	-	-	-	-	0.210422
	t-stat	-	-	-	-	-	*0.65*
	Coeff-VARCOND	40.6147	40.4494	89.7696	89.5557	64.3093	64.3376
	t-stat	*5.30*	*5.28*	*10.71*	*10.69*	*7.81*	*7.82*
	Coeff-VARCONDSECT	184.927	184.978	221.79	222.027	229.862	229.854
	t-stat	*65.72*	*65.73*	*71.99*	*72.02*	*77.08*	*77.08*
	R²	0.3848	0.3848	0.2657	0.2658	0.2881	0.2881

Figure 1

Daily number of news releases from January 1999 to December 2003

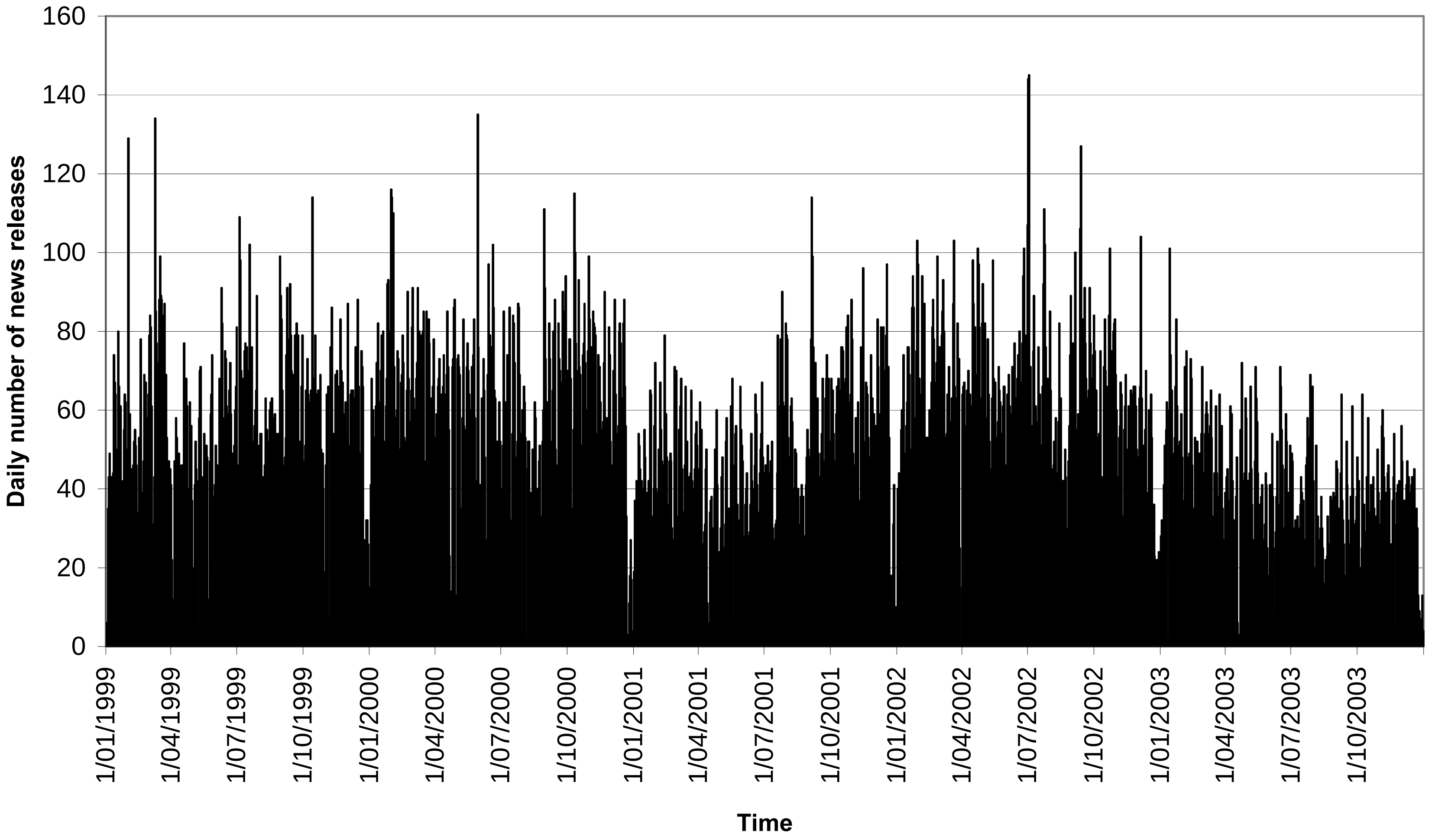


Figure 2

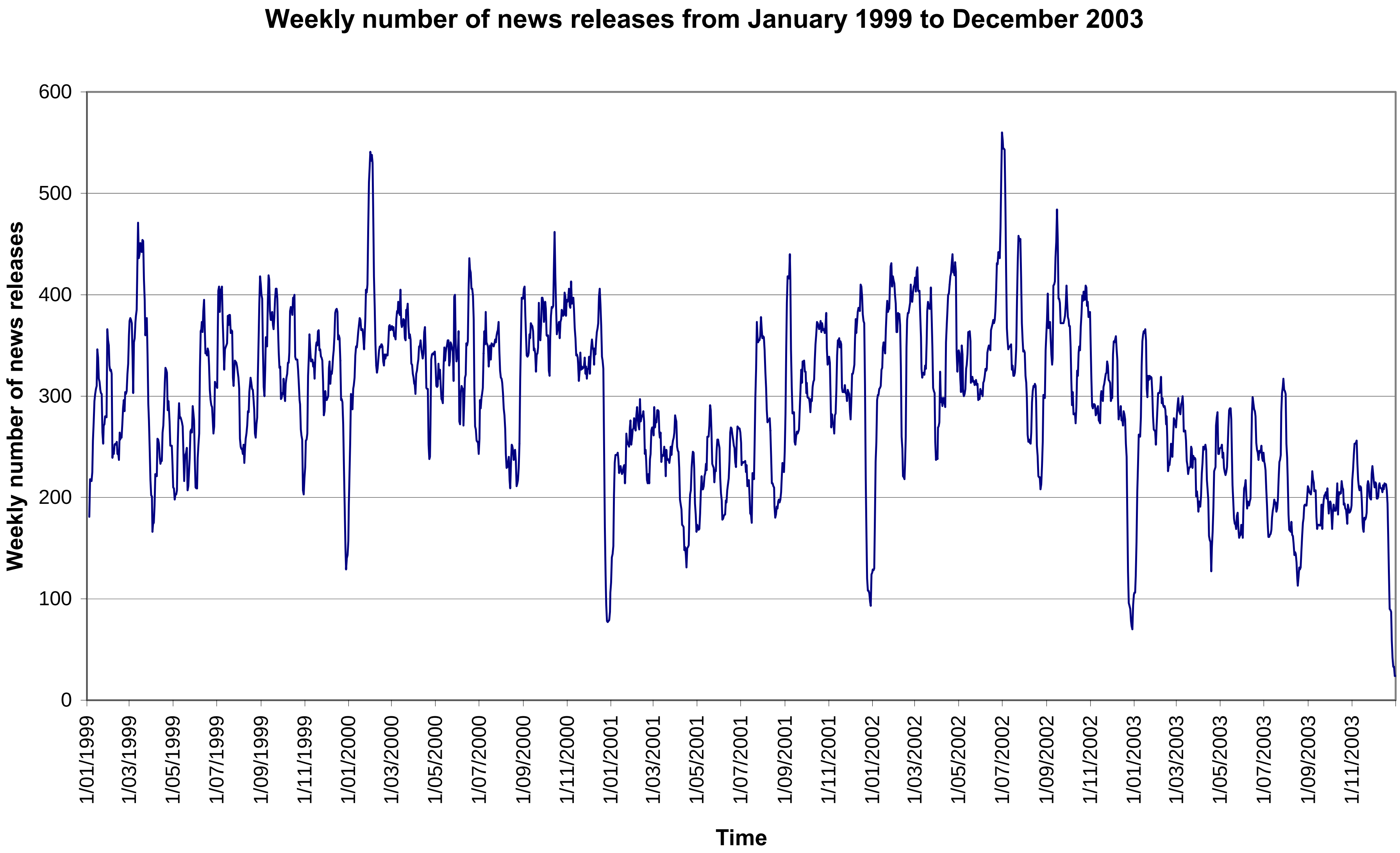

Weekly number of news releases from January 1999 to December 2003
Weekly number of news releases
600
500
400
300
200
100
0
1/01/1999
1/03/1999
1/05/1999
1/07/1999
1/09/1999
1/11/1999
1/01/2000
1/03/2000
1/05/2000
1/07/2000
1/09/2000
1/11/2000
1/01/2001
1/03/2001
1/05/2001
1/07/2001
1/09/2001
1/11/2001
1/01/2002
1/03/2002
1/05/2002
1/07/2002
1/09/2002
1/11/2002
1/01/2003
1/03/2003
1/05/2003
1/07/2003
1/09/2003
1/11/2003
Time

Figure 3

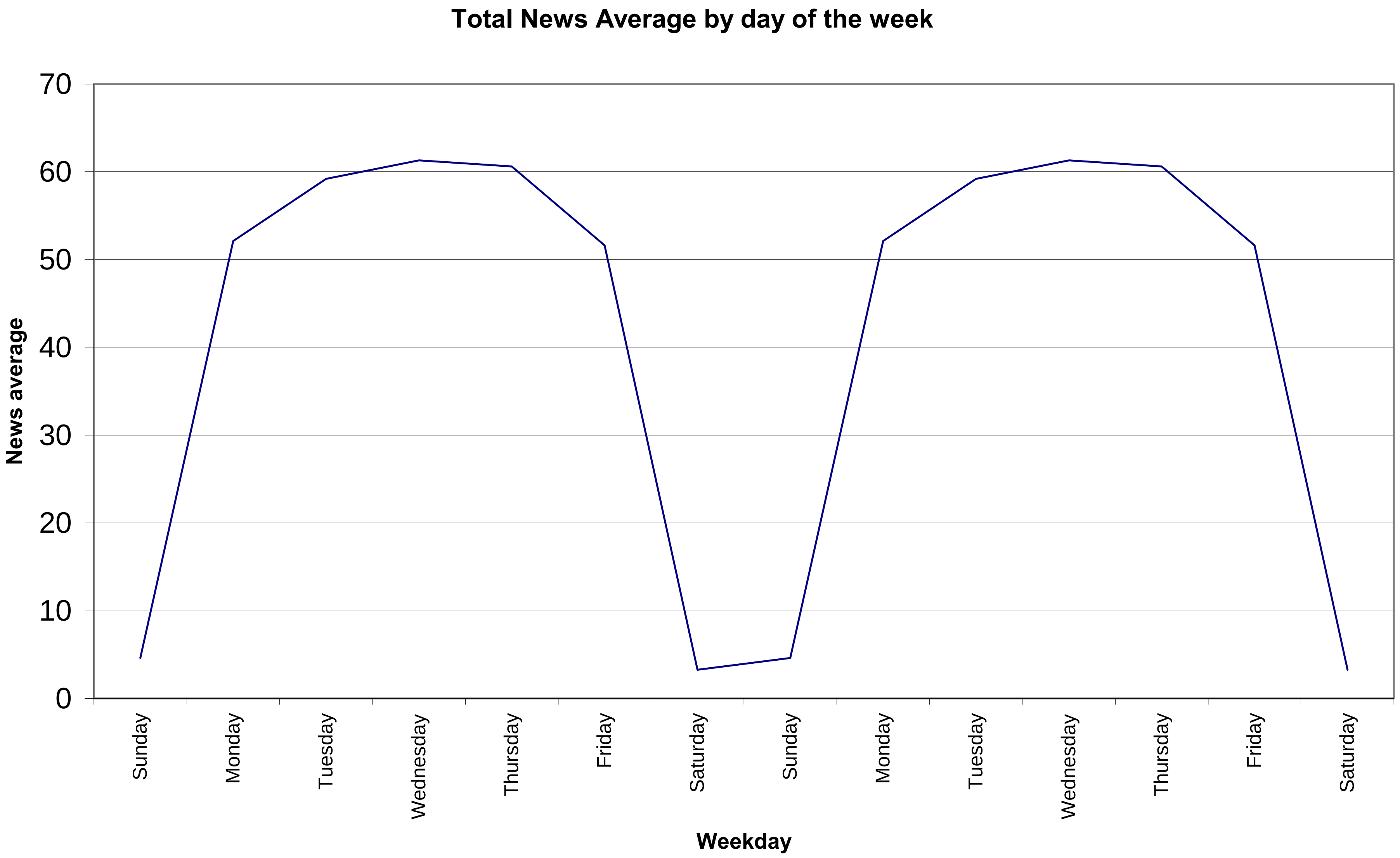

Total News Average by day of the week
News average
70
60
50
40
30
20
10
0
Sunday
Monday
Tuesday
Wednesday
Thursday
Friday
Saturday
Sunday
Monday
Tuesday
Wednesday
Thursday
Friday
Saturday
Weekday

Figure 4

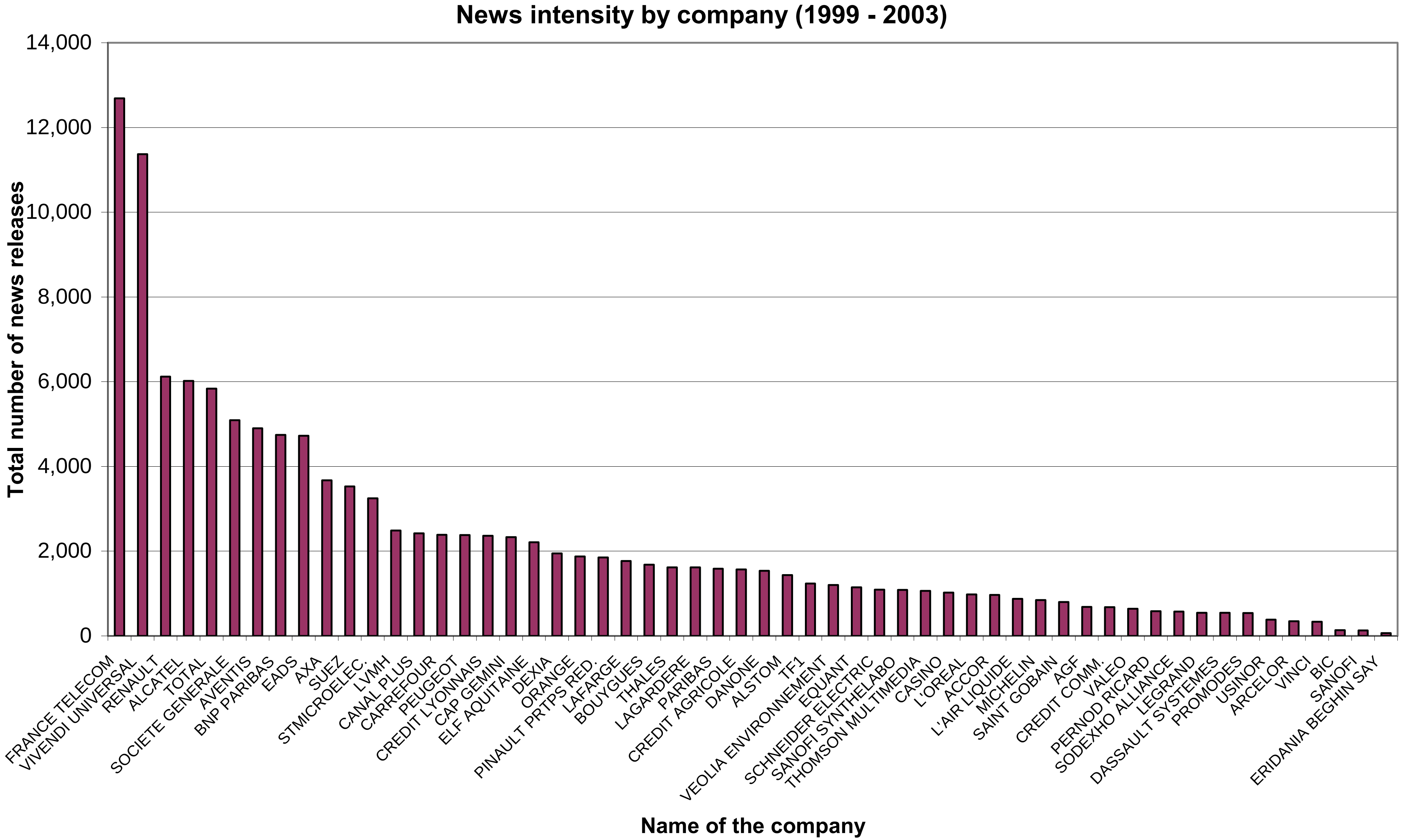
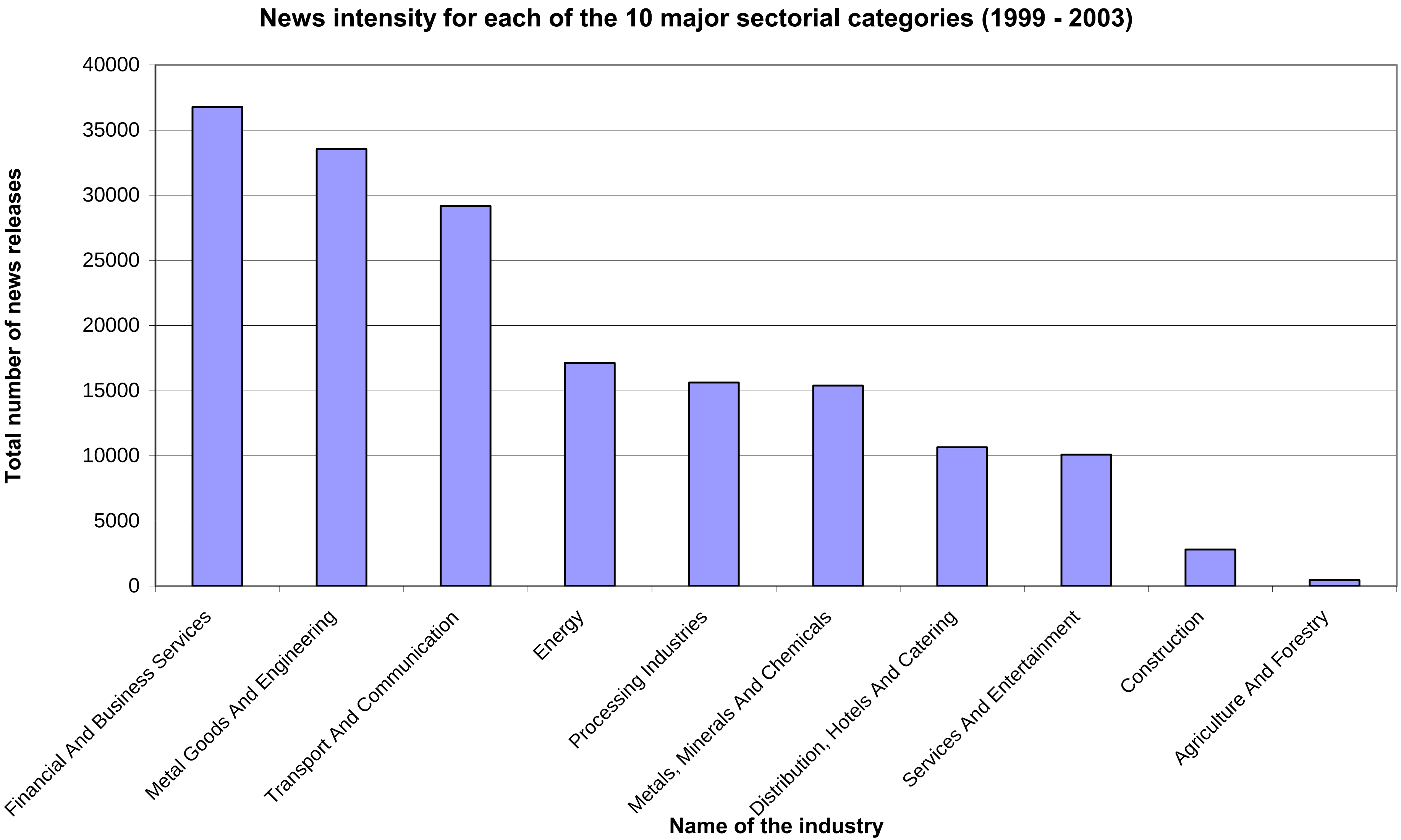

News intensity by company (1999 - 2003)
Total number of news releases
14,000
12,000
10,000
8,000
6,000
4,000
2,000
0
FRANCE TELECOM
VIVENDI UNIVERSAL
RENAULT
ALCATEL
TOTAL
SOCIETE GENERALE
AVENTIS
BNP PARIBAS
EADS
AXA
SUEZ
STMICROELEC.
LVMH
CANAL PLUS
CARREFOUR
PEUGEOT
CREDIT LYONNAIS
CAP GEMINI
ELF AQUITAINE
DEXIA
ORANGE
PINAULT PRTPS RED.
LAFARGE
BOUYGUES
THALES
LAGARDERE
PARIBAS
CREDIT AGRICOLE
DANONE
ALSTOM
TF1
VEOLIA ENVIRONNEMENT
EQUANT
SCHNEIDER ELECTRIC
SANOFI SYNTHELABO
THOMSON MULTIMEDIA
CASINO
L'OREAL
ACCOR
L'AIR LIQUIDE
MICHELIN
SAINT GOBAIN
AGF
CREDIT COMM.
VALEO
PERNOD RICARD
SODEXHO ALLIANCE
LEGRAND
DASSAULT SYSTEMES
PROMODES
USINOR
ARCELOR
VINCI
BIC
SANOFI
ERIDANIA BEGHIN SAY
Name of the company

Figure 5



News intensity for each of the 10 major sectorial categories (1999 - 2003)
Total number of news releases
40000
35000
30000
25000
20000
15000
10000
5000
0
Financial And Business Services
Metal Goods And Engineering
Transport And Communication
Energy
Processing Industries
Metals, Minerals And Chemicals
Distribution, Hotels And Catering
Services And Entertainment
Construction
Agriculture And Forestry
Name of the industry

Figure 6

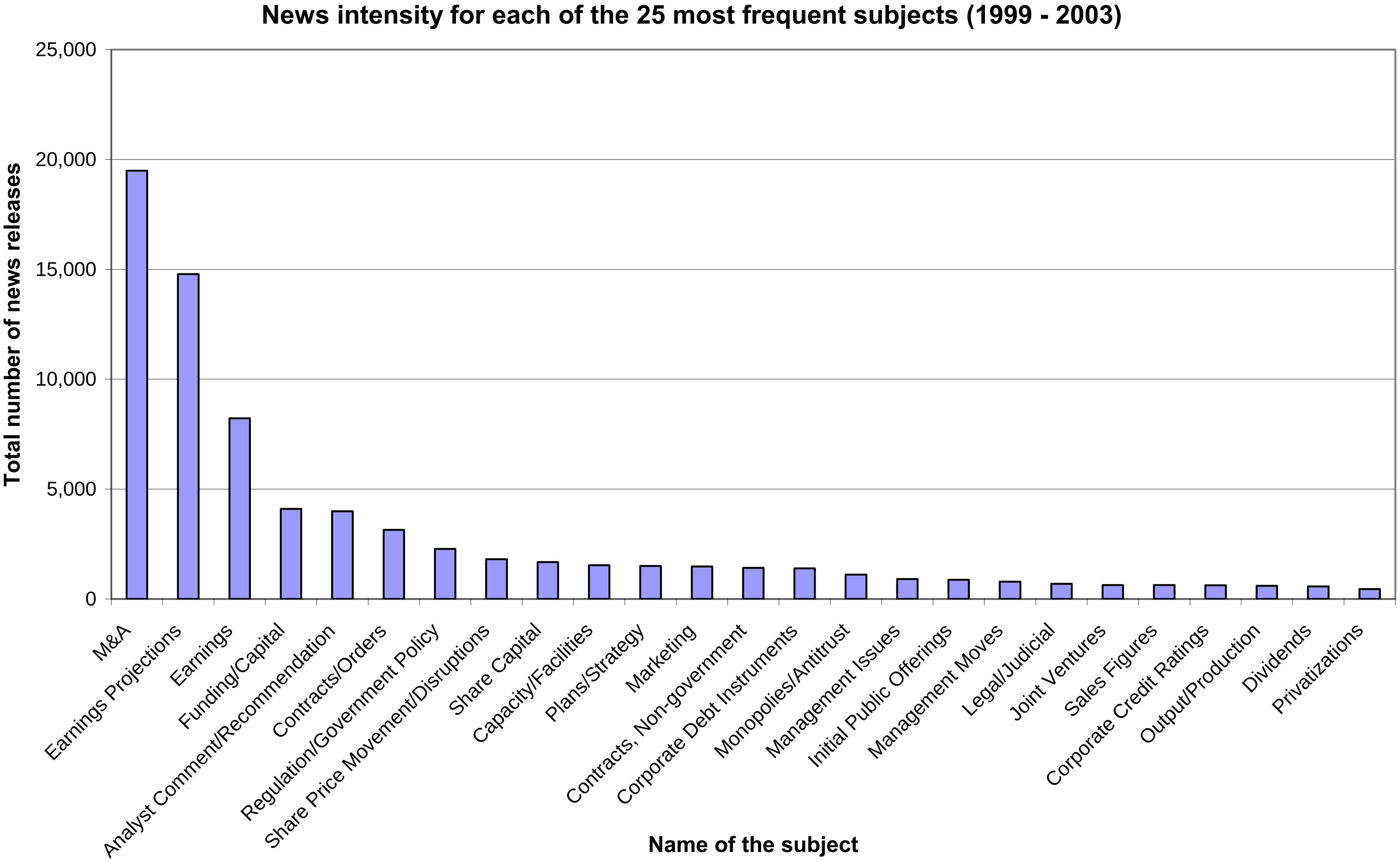

News intensity for each of the 25 most frequent subjects (1999 - 2003)
Total number of news releases
25,000
20,000
15,000
10,000
5,000
0
M&A
Earnings Projections
Earnings
Funding/Capital
Analyst Comment/Recommendation
Contracts/Orders
Regulation/Government Policy
Share Price Movement/Disruptions
Share Capital
Capacity/Facilities
Plans/Strategy
Marketing
Contracts, Non-government
Corporate Debt Instruments
Monopolies/Antitrust
Management Issues
Initial Public Offerings
Management Moves
Legal/Judicial
Joint Ventures
Sales Figures
Corporate Credit Ratings
Output/Production
Dividends
Privatizations
Name of the subject

Figure 7

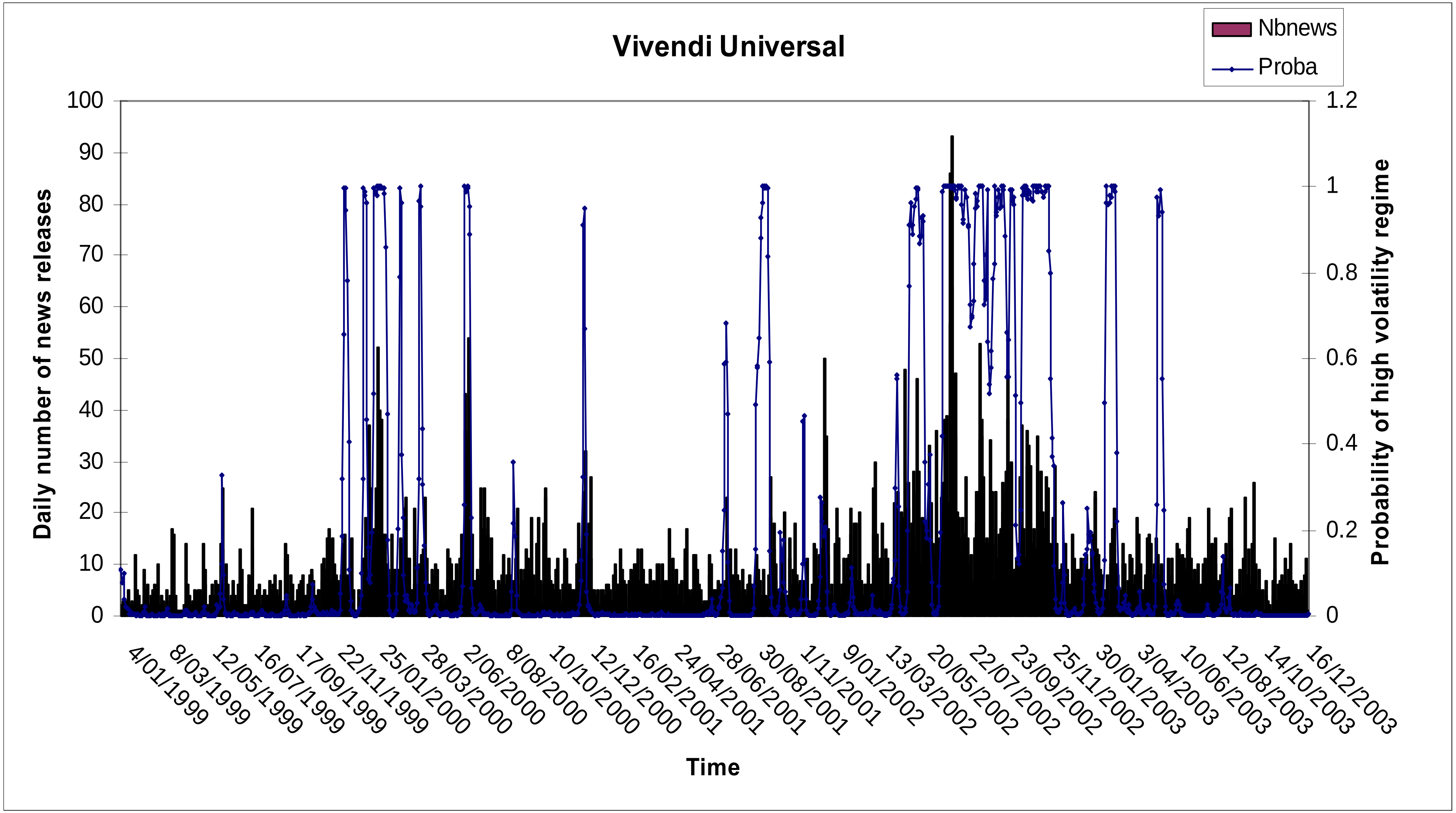
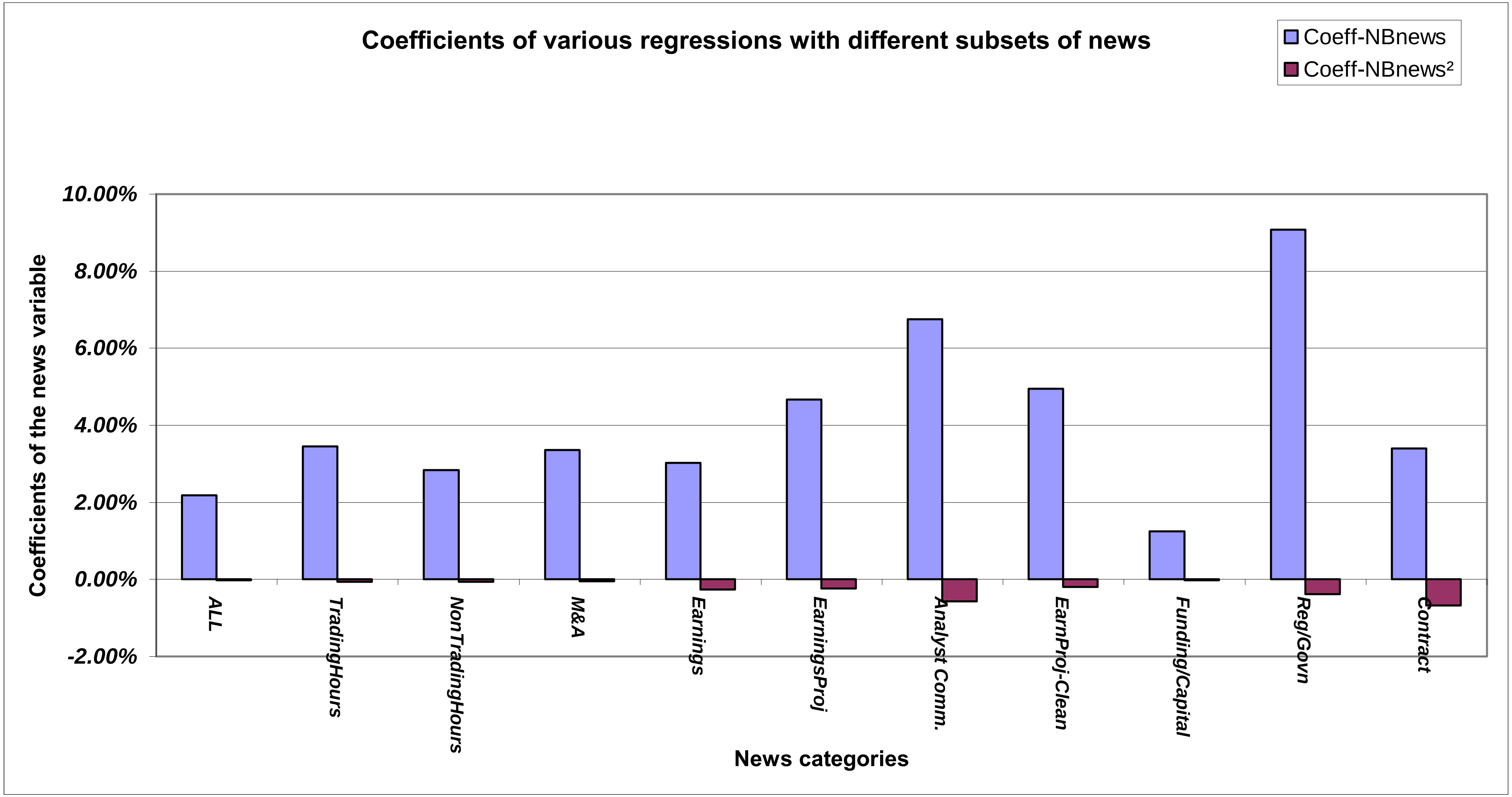

Vivendi Universal
Nbnews
Proba
Daily number of news releases
100
90
80
70
60
50
40
30
20
10
0
Probability of high volatility regime
1.2
1
0.8
0.6
0.4
0.2
0
4/01/1999
8/03/1999
12/05/1999
16/07/1999
17/09/1999
22/11/1999
25/01/2000
28/03/2000
2/06/2000
8/08/2000
10/10/2000
12/12/2000
16/02/2001
24/04/2001
28/06/2001
30/08/2001
1/11/2001
9/01/2002
13/03/2002
20/05/2002
22/07/2002
23/09/2002
25/11/2002
30/01/2003
3/04/2003
10/06/2003
12/08/2003
14/10/2003
16/12/2003
Time

Figure 8


Coefficients of various regressions with different subsets of news
Coeff-NBnews
Coeff-NBnews²
10.00%
8.00%
6.00%
4.00%
2.00%
0.00%
-2.00%
Coefficients of the news variable
ALL
TradingHours
NonTradingHours
M&A
Earnings
EarningsProj
Analyst Comm.
EarnProj-Clean
Funding/Capital
Reg/Govn
Contract
News categories

APPENDIX A

Proportion data regression

The dependent variable is the proportion (P_i) of the ni individuals. The regression analysis of P_i, as shown in Greene (2000, p. 835), raises a concern of heteroscedasticity. The observed P_i is an estimate of the population quantity, $\pi_i = F(\beta X_i)$. If we treat this problem as sampling from Bernoulli population, then we have:

$$P_i = F(\beta X_i) + \varepsilon_i = \pi_i + \varepsilon_i \qquad \text{(A2.1)}$$

where:

$$E[\varepsilon_i] = 0 \, , \; Var[\varepsilon_i] = \frac{\pi_i(1-\pi_i)}{n_i} \qquad \text{(A2.2)}$$

This heteroscedastic regression format suggests that the parameters could be estimated by a nonlinear weighted least squares regression. But the author proposes a simpler way to proceed. Since the function $F(\beta X_i)$ is strictly monotonic, it has an inverse.

$$F^{-1}(P_i) = F^{-1}(\pi_i + \varepsilon_i) \approx \beta X_i + \frac{\varepsilon}{\int_i} \qquad \text{(A2.3)}$$

This equation produces a heteroscedastic linear regression:

$$F^{-1}(P_i) = Z_i = \beta X_i + u_i \qquad \text{(A2.4)}$$

where:

$$E[u_i] = 0 \, , \; Var[u_i] = \frac{F_i(1-F_i)}{n_i \int_i^2} \qquad \text{(A2.5)}$$

The inverse function for the logistic model is easy to obtain. If

$$\pi_i = \frac{\exp(\beta X_i)}{1+\exp(\beta X_i)} \qquad \text{(A2.6)}$$

then:

$$Ln\left(\frac{\pi_i}{1-\pi_i}\right) = \beta X_i \qquad \text{(A2.7)}$$

Weighted least squares regression produced the minimum χ^2 estimator of β. Since the weights are function of the unknown parameters, a two step procedure is called for. Simple least squares at the first step produces a consistent but inefficient estimate. Then the weights for the logit model based on the first step estimates are then:

$$W_i = n_i \Lambda_i (1-\Lambda_i) \text{ with } \Lambda_i = \frac{\exp(\beta X_i)}{1+\exp(\beta Xi)} \tag{A2.8}$$

and can be used for weighted least squares in the second step procedure.