NHH

Norwegian School of Economics

Bergen, Fall, 2019

High-Frequency Market Reactions to Unscheduled Stock-Specific News

An Empirical Analysis of the Intraday Market Dynamics at the Oslo Stock Exchange

Philip Baltzersen Holm & Jørgen Rødde

Supervisor: Maximilian Rohrer

Master thesis, Economics and Business Administration,

Financial Economics

NORWEGIAN SCHOOL OF ECONOMICS

Abstract

We use pre-processed news data combined with high-frequency stock data from the Oslo Stock Exchange to test the following hypotheses: (1) the sentiment of news articles can predict the direction of intraday abnormal returns; and (2) intraday volatility and trading activity increase around the arrival of news articles. First, we find that abnormal returns are significantly negative at 17 basis points 90 minutes after a negative news release. In contrast, we cannot establish a significant relationship between abnormal returns and news with a positive or neutral sentiment. Second, by using a high-frequency vector autoregressive model, we find that: (1) volatility increase on average 0.47 standard deviations ten minutes before a news arrival; and (2) money value traded increase by 0.48 and 0.47 standard deviations five and ten minutes before news arrivals. Thus, our results suggest that negative news articles affect the abnormal returns more than positive news articles and that unscheduled news affects the intraday volatility and trading activity at the Oslo Stock Exchange.

Acknowledgements

We want to thank our supervisor Maximilian Rohrer for his valuable support and crucial comments throughout the process. Further, we are grateful to Joakim Blix Prestmo for providing us with continuous feedback. Finally, we want to thank Strise for providing us with the data that has made this thesis possible.

Contents

List of Tables

List of Figures

1 Introduction

There is a considerable amount of literature analyzing the role of media in the financial market. However, there are only a couple of studies examining high-frequency market reactions to unscheduled news.[1] One reason why few have studied this area of the literature is that the news is very noisy and hard to quantify. The continuous flow of information from mass media makes it challenging to determine which news is relevant. Due to advances in machine learning and natural language processing, it is now possible to quantify this constant flow of information. By using unique pre-processed news data combined with intraday stock data, we link unscheduled stock-specific news to high-frequency market dynamics at the Oslo Stock Exchange. Thus, we aim to extend this limited area of the financial literature regarding how unscheduled news can explain high-frequency movements in volatility and trading activity.

Our first hypothesis states that the sentiment of stock-specific news can predict the direction of intraday abnormal returns. To test this hypothesis, we perform an intraday event study. By aggregating the stock data into 40-second intervals and separating the news articles into positive, negative, and neutral, we compute the average cumulative abnormal return across every news article. The event window covers 60 minutes before and 90 minutes after the arrival of a news article. We find that news with negative sentiment has significant negative abnormal returns from 60 minutes to 30 minutes before the news release, before reverting until the time of arrival. There are also significant negative abnormal returns from the initial arrival of news to 90 minutes after disclosure at 17 basis points. Conversely, we do not find significant abnormal returns for news with a positive or neutral sentiment.

Secondly, we hypothesize that volatility and trading activity increase around the arrival of news. To get an overall view of the market dynamics around news arrivals, we first investigate how volatility and trading activity respond on average. This analysis indicates that both increase when a news article arrives. Further, to take into consideration the autocorrelation and cross-correlation in the data, we employ a high-frequency vector autoregressive (VAR) model with 40-second and 5-minute intervals for each firm in the OBX Index. We add dummy variables to represent the arrival of news and then compute the average model estimates from the 25 VAR models. Our results show that there is no significant reaction in any of the variables to the news with 40-second intervals. However, with 5-minute intervals, we find that money value

[1] To our best knowledge, only Groß-Klußmann and Hautsch (2011) and Smales (2014) have studied this topic.

traded on average increase by 0.48 and 0.47 standard deviations five and ten minutes before the initial release of news articles, both significant at the 10% level. Moreover, we find a significant increase in volatility of 0.47 standard deviations, ten minutes before the news disclosure at the 10% level. Additionally, there are significant reactions in almost every variable in the individual VAR models for each firm, both with 40-second and 5-minute intervals. However, the effects are diffuse across the companies.

To validate our results from the 5-minute VAR model, we run three additional models. First, we randomly assign the news articles to firms and run the model with the same specifications, and find that news arrival causes no significant market reactions. Second, we control for weekday effects by including weekday-dummies in the model. The results are the same as the original model, which means that weekday effects do not drive our results. Lastly, we control for intraday seasonality patterns by including dummy variables that represent each 5-minute interval in a trading days. Since the same news dummies are significant, it suggests that our results also are robust to intraday seasonality patterns.

To summarize, our results from the event study are consistent with the area of the literature that argues information with a negative sentiment impacts the stock market to a higher degree than information with a positive sentiment. However, the abnormal return patterns are not consistent with the literature, as there is no clear trend for positive or negative news. This may be due to differences in the sentiment indicators and does not indicate that the sentiment of a news article cannot predict the direction of intraday abnormal returns. Rather, it highlights the difficulty of correctly identifying the tone of textual information. Furthermore, the results from the VAR models support the current literature regarding increases in volatility and trading activity around arrivals of unscheduled news. The fact that we find the same relationship between news arrivals and high-frequency market reactions as Groß-Klußmann and Hautsch (2011) and Smales (2014) implies that the relationship is universal across stock markets. It is especially noteworthy that we achieve the same results considering that all three studies use different news engines and analyze different stock markets.

1.1 Theoretical background and hypothesis development

Fama (1970) proposes that, according to the efficient market hypothesis, it should not be possible to earn risk-adjusted abnormal returns trading on publicly available information. He argues that a firm's stock price should quickly incorporate new information. Conversely, Roll (1988) was not successful in finding a significant relationship between news and stock returns (see also Schwert (1981) and Cutler, Poterba, and Summers (1989)). Boudoukh, Feldman, Kogan, and Richardson (2013) propose an explanation of why the literature had a hard time finding a causal relationship between the news and the stock market. They suggest that the previous literature has merely been doing a poor job at correctly quantifying news and found considerably more evidence of a relationship between news and stock prices.

Regardless of the relationship between news and market reactions, old information should not affect the market (Fama, 1970). However, Tetlock (2011) finds that stale news, which is defined by its textual similarity to previous news stories, affects the market. He also argues that individual investors trade more on news that is stale, which leads to temporary movements in a firm's stock prices. Huberman and Regev (2001) analyze a prominent example when the stock market reacted to stale news. A large newspaper published a front-page story about a promising anti-cancer drug, which had already been published in a scientific journal five months earlier. The stock price of the firm that developed the drug increased by over 600% in one day. One argument for this phenomenon is that media can mitigate asymmetrical information between investors (Fang & Peress, 2009). Publicly available information could be known only to a portion of investors due to attention bias, which arises due to the limited amount of information a person can consider (Bachmann, Enrico, & Hens, 2018). Hence, until the media writes about a news story, some investors are not able to consider that information. Thus, the financial media broadcast stale news to a subgroup of investors that unintentionally cause mispricing in the short-run (Tetlock, 2011).

As the literature suggests, a firm's stock price reacts to news, even when it is stale. However, to be able to predict which direction the stock price will move, textual analysis is required to distinguish between news with positive and negative information. Tetlock (2007) was one of the first to use automated textual analysis on news articles. He finds that negative words are associated with lower prices in a short period before reverting. Later, Tetlock, Saar-Tsechansky, and Macskassy (2008) find that the fraction of negative words in stock-specific news stories forecasts low earnings. They suggest that words within news stories capture firms' fundamentals that are hard to interpret. Moreover, several other studies show how textual information can predict returns both for individual securities and the aggregate market level (see, e.g., Chen, De, Hu, and Hwang (2014), Dougal, Engelberg, García, and Parsons (2012), García (2013), and Heston and Sinha (2017)). Based on the vast amount of literature that use textual information to predict stock returns, we expect the stock price to increase around the arrival of news with a positive sentiment and decrease around the arrival of news with a negative sentiment. For news with a neutral sentiment, we do not expect the stock price to react. Thus, we propose the first hypothesis:

1. The sentiment of news articles can predict the direction of intraday abnormal returns.

To address how the news affects trading activity, we first examine some biases that could explain why investors trade more when information arrives. Tversky and Kahneman (1973) find that people evaluate the probability of events by availability, which leads to systematic biases. Barber and Odean (2008) use this bias to find that individual investors buy stocks that get much attention in the news. Moreover, Odean (1999) finds that trading volume increases when individual investors are overconfident. These biases could help explain why Mitchell and Mulherin (1994) and Berry and Howe (1994) find that the number of news releases and trading volume are positively correlated. These findings are further strengthened by Peress (2014), who investigates the causal impact of media on trading activity. By using strikes in national newspapers, he finds that trading volume falls by 12% on days in which a strike happens. This is also in line with Engelberg and Parsons (2011), who find that local media coverage increased the local trading activity of individual investors by 48%.

With regard to how volatility responds to news arrivals, Kalev, Liu, Pham, and Jarnecic (2004) find that the arrival rate of news significantly affects the volatility of stock returns. This is further supported by Peress (2014), who finds that intraday volatility is reduced by 7% in the event of a newspaper strike. One reason why volatility increases when information arrives might be that traders interpret the same information differently (M. Harris & Raviv, 1993). This can explain why Antweiler and Frank (2004) are able to predict market volatility using the information on stock message boards. Collectively, these findings support the mixture of distribution hypothesis, which states that the volatility at a given interval is proportional to the arrival rate of information (L. Harris, 1987). It also states that changes in price and volume are correlated since they are both driven by the same information arrival processes. Thus, based on the literature above, there is a reason to believe that both volatility and trading activity increase around news arrival. We propose the following hypothesis:

2. Intraday volatility and trading activity increase around the arrival of news stories from the financial media.

1.2 Literature review

Our study contributes to several different fields of the financial literature. First, we relate to a large number of studies using textual information to explain and predict stock price movements. The literature suggests that it is possible to predict stock prices to some extent, and that information with negative sentiment has the biggest effect on the market (see, e.g., Tetlock (2007), Tetlock et al. (2008), and García (2013)). Second, we relate to the literature that analyze intraday market reactions to scheduled news, such as earnings announcements (see, e.g., Patell and Wolfson (1984) and Lee (1992)). A big advantage of using scheduled news is that it is easily identifiable and relevant to the respective firm. Lastly, we also relate to the area of literature that analyzes market reactions to unscheduled news. However, it has been challenging to filter out relevant news due to the excess noise. Thus, the news flow usually has been aggregated to a daily level (see, e.g., Berry and Howe (1994) and Mitchell and Mulherin (1994)). As discussed, there has not been a considerable amount of research examining high-frequency market reactions to unscheduled news. However, Groß-Klußmann and Hautsch (2011) and Smales (2014) analyze this topic, which is why most of our framework is based on those studies.

Groß-Klußmann and Hautsch (2011) analyze the intraday market reactions to stock-specific unscheduled news by using data from the Reuters NewsScope Sentiment Engine. Similar to us, their dataset contains machine-generated sentiment and relevance filters based on linguistic pattern recognition. By aggregating the return data into 20-second intervals, they study high-frequency market reactions to 29 487 news articles over one year. In relation, our sample consists of 908 news articles over 12 weeks. Their first hypothesis is that there are theory-consistent market reactions in returns, volatility, and liquidity. They find that, on average, volatility and trading activity increase significantly around the arrival of news. This is consistent with our estimates of the average reaction on volatility and trading activity. When controlling for autocorrelation and cross-correlation in a VAR model, Groß-Klußmann and Hautsch (2011) only find a significant reaction in volatility and money value traded. Likewise, we find a significant reaction in the same variables. However, our results indicate that the reaction appears five and ten minutes before the news disclosure, instead of exactly when the news arrives.

Their second hypothesis is whether trading based on an intraday news flow is profitable. They use an event study to analyze the abnormal returns for positive, negative, and neutral sentiment labeled news articles. Although their news engine seems to identify the direction of abnormal returns correctly, the majority of the reaction appeared before news-disclosure. Hence, they fail to find a profitable trading strategy after controlling for trading costs. In contrast, we find some evidence that negative news leads to negative abnormal returns before and after the news disclosure. However, we cannot predict the correct direction of abnormal returns for positive news articles.

Similarly, Smales (2014) also bases his framework on Groß-Klußmann and Hautsch (2011). He explores the market reactions at the Australian Securities Exchange to a pre-processed news flow based on the news engine from Ravenpack. By analyzing 484 400 news headlines over ten years, his results is consistent with Groß-Klußmann and Hautsch (2011) as he finds that volatility and money value traded increases significantly around news arrivals. Additionally, he finds that unscheduled news has a significant impact on average trade size, absolute order imbalance, and bid-ask spreads. In contrast, we fail to find a significant relationship between the intraday news flow and average trade size.

There is not much literature regarding the Norwegian stock market and intraday market reactions. However, Larsen and Thorsrud (2017) analyze the content of the Norwegian business paper Dagens Næringsliv by classifying each article into topics and sentiment. Using stock prices on a daily level and around 250 000 news articles, they find that that the content of the news articles significantly predicts daily returns. Moreover, they find that a zero-cost trading strategy based on the sentiment of the news articles leads to annualized risk-adjusted returns of 20%. In contrast, we investigate returns at a higher frequency, and we use several news providers, both Norwegian and international. However, our sample period is considerably smaller. Consequently, by analyzing how financial media affects the Norwegian stock market at higher frequencies, we hope to contribute to the current literature in a global context.

2 Data

In this section, we describe our data sources, the data processing, and assumptions we make regarding our data. Our data span 12 weeks, from 26.08.19 to 15.11.19, and consists of two parts: stock data from the Oslo Stock Exchange and news data from Strise. The combined data set will be used for empirical analysis.

The stock data consists of every single trade (timestamp, price, volume, and type) for the 25 companies listed in the OBX Index. There are mainly two different types of trades: automatic trades and off-book trades. The latter are trades that are subject to conditions other than the market price, such as non-ordinary trades, OTC trades, and trades related to derivatives. By examining the off-book trades, we find that they sometimes deviate drastically from nearby trades. As shown in Figure 10 in the Appendix, we observe that some of the off-book trades cause outliers. By excluding off-book trades from our analysis, we avoid the extreme return observations. The stock price without off-book trades is visualized to the right in Figure 10.

Although the companies in the OBX Index are the most liquid stocks at the Oslo Stock Exchange, there are still frequent periods with no trades. Analyzing the data with 1-second intervals would not only be impractical due to computational limitations, but there would also be many numbers of intervals without any trade. This can lead to distributional misspecifications, which would cause inconsistent parameter estimates (Hautsch, Malec, & Schienle, 2013). Consequently, we aggregate the data into 40-second intervals and 5-minute intervals. The former has 40 560 intervals per company and 46% observations on average with no trade. The latter has 5 460 intervals per company and only 5% observations on average with no trade. Moreover, intraday price data is contaminated by market noise (Duan, Härdle, & Gentle, 2012). Hence, we follow Barndorff-Nielsen and Shephard (2002) and use realized volatility as our measure of intraday volatility. Furthermore, to capture trading activity, we compute the money value traded, the average trade size, and the number of trades. After processing the stock data, we have the following variables

1. Return: Calculated with the average price from each interval. For intervals without any trade, we use the price from the previous interval with a trade.

2. Volatility: Realized volatility, calculated as $RV_\tau = \sqrt{\sum_{i=1}^{m} r_i^2}$, where m is the number of trades in interval τ and r_i is the return from one trade to another.

3. Money value traded: The cumulative sum of price multiplied with volume for every trade in an interval.

4. Average trade size: The total number of shares traded divided by the total number of trades in an interval.

5. Number of trades: The total number of trades in an interval.

The news data is based on unstructured data from over 100 000 different publishers. We filter out the publishers that are finance-related and have published at least one story about a firm in the OBX Index in the sample period. To filter out noise, we only examine news articles that are labeled as relevant. This means that the story is about the specific firm, and not just mentions it. Afterward, 23 unique financial media sources remain, and there is a total of 906 news articles in the sample period. Moreover, each article contains a timestamp with precision by a second and a variable to indicate its sentiment (positive, negative, and neutral). In Table 1, four examples are shown from the news data.

Table 1: Examples from news data.

Entity name	Story published	Publisher	Story title	Relevant	Entity summary	Sentiment
Equinor ASA	2019-09-13 12:20:27	uk.reuters.com	UPDATE 2-Fire breaks out on tanker at Norway's Sture oil terminal	FALSE	OSLO, Sept 13 - A fire has broken out in the engine room of an oil tanker during loading at Equinor's Sture oil export terminal on Norway's west coast, local police and the oil firm said on Friday.	Neutral
Equinor ASA	2019-08-27 16:09:42	uk.reuters.com	Sri Lanka to start oil production in 2023 Total, Equinor to study potential	TRUE	Equinor will bear 30% of the project cost, its Vice President for exploration Janne Rui said, without specifying details.	Neutral
Equinor ASA	2019-11-06 12:45:39	sg.finance.yahoo.com	Equinor makes Norway oil, gas find of up to 100 million barrels	TRUE	OSLO - Norwegian energy major Equinor and its partners have found oil and gas at the Echino South prospect near the Fram field in the North Sea, the company said on Wednesday.	Positive
Equinor ASA	2019-09-03 14:02:22	aksjelive.e24.no	Dansk pensjonsfond selger seg ut av ti oljeselskaper	TRUE	Martin Hagh Høgseth Det danske pensjonsfondet MP Pension har besluttet å selge alle sine aksjer i ti av verdens største oljeselskaper, blant disse også Equinor.	Negative

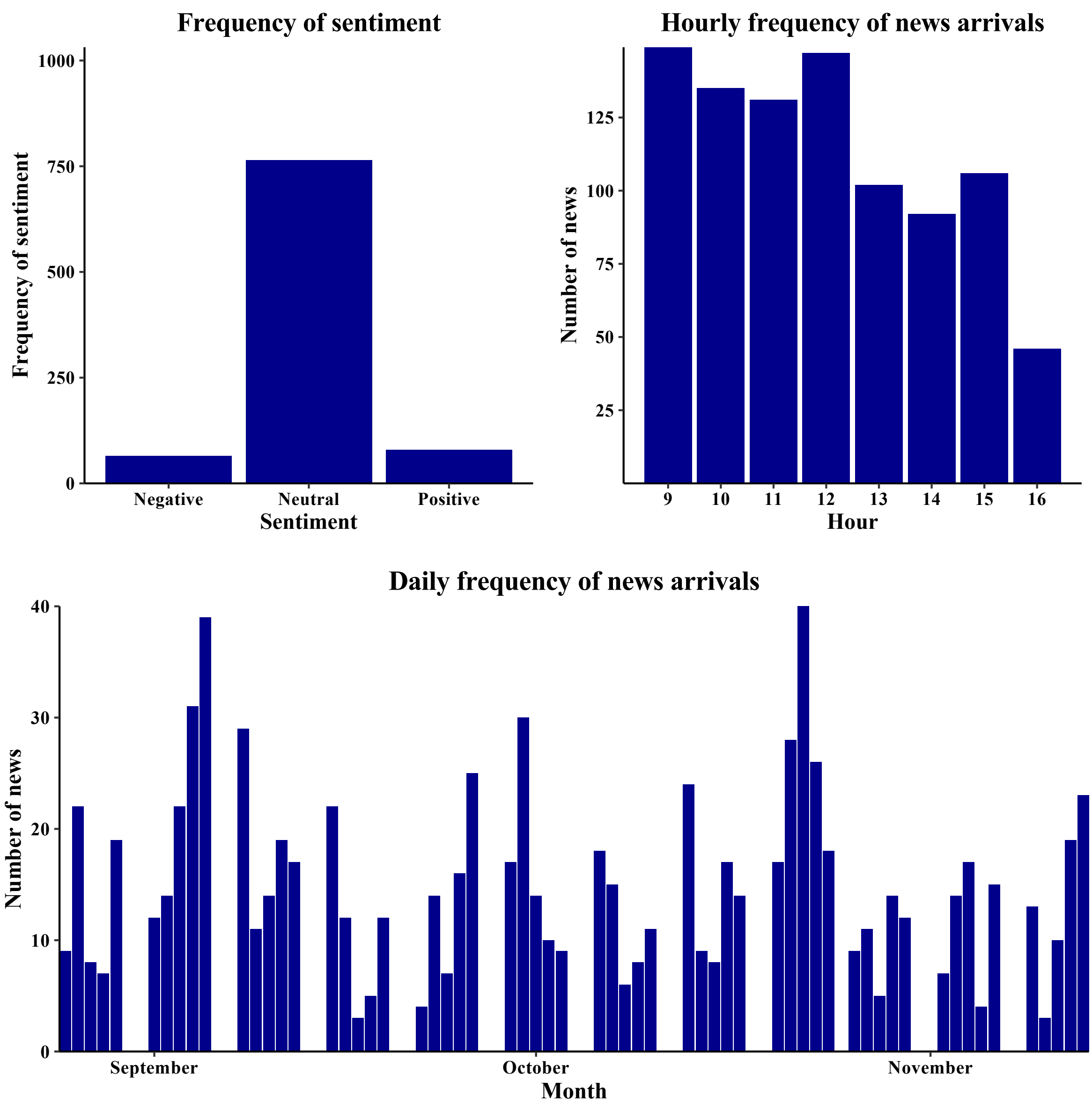

Figure 1: From the top left corner, the first figure shows the frequency of news articles with a different sentiment. The second figure illustrates the frequency of news arrivals based on the hour of the day. The last hour only shows the number of news arrivals for 30 minutes since the stock exchange closes at 16:30. The figure at the bottom shows the frequency of news arrivals at a daily level.

The distribution of articles based on its sentiment is shown in the upper left corner in Figure 1. There are 764 neutral, 79 negative, and 65 positive articles. In other words, 84% of our news articles are classified as neutral, and consequently, the news data is unbalanced. Moreover, as our objective is to identify intraday market reactions, we have filtered out news stories that were published outside the opening hours (09:00 - 16:30) of the Oslo Stock Exchange. The largest number of news arrives at the beginning of the trading day, as shown in the upper right corner in Figure 1. There is a decreasing trend throughout the day, with spikes at 12 and 15. The low amount of articles in the last hour is due to the fact that the Oslo Stock Exchange closes at 16:30. Lastly, the bottom of Figure 1 visualizes the daily frequency of news articles. The first week in September and the fourth week in October have the most articles, with 118 and 129, respectively. The average number per week is 75, and there is no clear trend in the data on a daily level. Furthermore, we remove the first and last 15 minutes of each trading day from the data set. By reducing the opening and close effects, we avoid observing large spikes in returns, volatility, and trading activity that are not necessarily caused by the news articles we are analyzing (Berry & Howe, 1994). Additionally, dates, where earnings reports were released, are filtered out. Thus, we reduce the effect of scheduled news on the results.

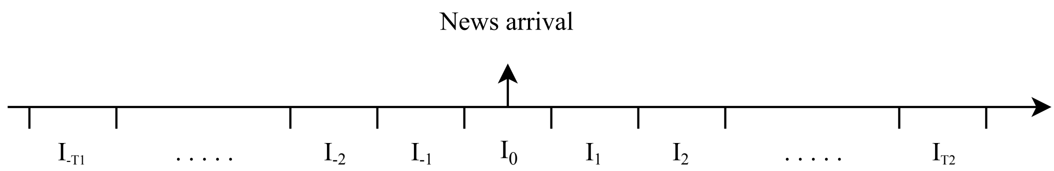

Figure 2: Intervals around news arrivals.

Figure 2 illustrates the intervals around news arrivals. The event window starts at $-T_1$, 90 intervals (60 minutes) before the news article is published, and ends at T_2, 135 intervals (90 minutes) after the news release. The data sets are merged by assigning each news article to an interval based on when it was published. When there are several news articles in a single interval, the sentiment indicator is summed, where positive counts as 1, negative counts as -1, and neutral is 0. Therefore, one positive $+1$ and one negative -1 in the same interval count as a neutral news article. Doing this, we do not consider that the news articles may have a different impact on market reactions, which can potentially weaken the reliability of our results. However, manually examining each article and give an impact score could lead to biases. Hence, we find that equally weighting the news articles are a reasonable assumption.

Descriptive statistics for each company are illustrated in Table 2. The firm with the most trades in our sample period is Equinor, with more than 400 000 trades. In contrast, the firm with the lowest number of trades is Elkem, with around 61 000 trades. Furthermore, there is a considerable difference between how many news articles there are for each firm. Equinor has the most with 294 news articles, and the least featured firm is Gjensidige Forsikring, with only three news articles.

Table 2: Descriptive statistics.

Ticker	MCAP %	Returns 26.08-09.11	Money value in MNOK	Average trade size	Nr. of trades	Nr. of news	Positive news	Negative news	Neutral news
AKER	1.8	15.4	2448.58	15	82969	14	3	1	10
AKERBP	4.8	20.45	7740.56	61	203939	60	2	5	53
BAKKA	1.75	11.56	2613.61	10	81385	23	4	1	18
BWO	0.61	46.31	3135.51	108	109581	20	1	2	17
DNB	13.18	13.58	9536.85	114	209124	42	6	0	36
DNO	0.55	-0.04	3874.71	656	184184	10	2	0	8
ELK	0.62	2.57	1128.69	147	61710	4	1	0	3
EQNR	28.92	18.57	22307.55	182	405411	294	23	25	246
GJF	4.29	-2.59	2532.00	34	112100	3	0	0	3
GOGL	0.36	-2.64	2280.17	104	85495	30	5	3	22
LSG	1.68	-8.24	3122.77	110	150558	4	1	0	3
MOWI	5.79	0.05	10529.01	82	214565	8	2	0	6
NEL	0.48	37.14	6229.16	1627	172478	7	1	0	6
NHY	3.36	27.16	8397.98	444	294508	65	3	8	54
NAS	0.28	33.51	5507.31	203	253900	58	2	4	52
ORK	4.49	7.18	4631.28	116	190392	11	0	1	10
PGS	0.26	55.11	2627.93	312	200094	7	1	0	6
SALM	2.22	-7.1	3918.86	19	159101	19	0	2	17
SCHA	1.41	-4.03	2045.23	18	116118	11	0	1	10
STB	1.54	27.76	3291.32	126	152262	15	0	2	13
SUBC	1.49	19.43	5034.9	97	247821	31	6	2	23
TEL	12.01	-8.21	9289.57	92	233898	102	5	5	92
TGS	1.51	25.64	3672.36	33	177600	5	0	0	5
TOM	1.91	-2.24	5238.03	37	173467	21	1	2	18
YAR	4.71	-5.86	6941.51	36	192596	44	10	1	33
Sum	100		138075.45		4465256	908	79	65	764

Note: Ticker is the ticker of the company, as indicated at the Oslo Stock Exchange. The second column shows each company's market capitalization as a percentage of the OBX Index. Returns are the total returns over the sample period. Money value is the total value traded in the sample period and is calculated by first multiplying the volume of each trade by its corresponding price, and then aggregating it across the entire sample period. Average trade size is the average number of shares per trade for each company. The number of trades shows the total number of trades over the sample period. The last four columns show how many relevant news articles there are in total and divided by sentiment.

3 Abnormal returns around news arrivals

3.1 Methodology

In order to test whether we can predict abnormal returns based on the sentiment of news articles, we use the event study framework outlined by Campbell, Lo, and MacKinlay (1996). An event study is used to measure the impact of an economic event on the market value of a firm (Campbell et al., 1996). The rationale behind the approach is, according to the efficient market hypothesis, that the impact of an event will be reflected immediately in the asset price (Fama, 1970). To measure the impact of an event, one must first estimate the normal return, which is the return that should take place if the event did not happen. We use the market model as our estimation model for the normal return because it reduces the abnormal return variance compared to the constant-mean-return model. In short, the market model relates the return of any company to the return of the market portfolio, and it removes the portion of the return that is related to the variation in the market's return. Theoretically, we could reduce the variance further by including additional risk factors, but the empirical gains have shown to be limited (Campbell et al., 1996).

As we do not have the intraday returns for the OBX Index, we have multiplied the returns for all stocks listed in the OBX Index with their corresponding weights (Second column in Table 2) and computed it as a proxy for the market portfolio's returns. The market return for interval τ is calculated by $R_{m,\tau} = \sum_{i=1}^{25} w_i r_{i,\tau}$, where w_i is the weight of company i in the index and $r_{i,\tau}$ is the return for company i at interval τ.

As discussed by Cont (2001), asset returns have significant autocorrelations for short intervals (less than 20 minutes). Thus, we use the Breusch-Godfrey test and find that returns for every company have significant autocorrelation at the 1% level (Wooldridge, 2013). To account for this, we include one lag of the stock return when we calculate the normal return, similar to Groß-Klußmann and Hautsch (2011). Hence, the normal return and the corresponding variance for the company i at interval τ is given by

$$R_{i,\tau} = \alpha_i + \beta_i R_{m,\tau} + \gamma_i R_{i,\tau-1} + \varepsilon_{i,\tau} \tag{1}$$

$$E[\varepsilon_{i,\tau}] = 0 \quad Var[\varepsilon_{i,\tau}] = \sigma^2_{\varepsilon_{i,\tau}}$$

where $R_{i,\tau}$ is the return for company i at interval τ, $R_{m,\tau}$ is the market return at interval τ, $R_{i,\tau-1}$ is the lagged return for company i, and $\varepsilon_{i,\tau}$ is the zero mean disturbance term. α_i, β_i, γ_i, and $\sigma^2_{\varepsilon_{i,\tau}}$ are the parameters of the market model. The normal return for company i over the entire estimation window (from T_0 to T_1) can be simplified by using matrices,

$$\boldsymbol{R}_i = \boldsymbol{X}_i \boldsymbol{\theta}_i + \boldsymbol{\varepsilon}_i \tag{2}$$

where $\boldsymbol{R}_i = [R_{i,T_0}, ..., R_{i,T_1}]'$ is a $(L_{est}\ x\ 1)$ vector of the estimation-window returns. L_{est} is the number of intervals in the estimation window for company i. $\boldsymbol{X}_i = [\boldsymbol{j}\ \boldsymbol{R}_m\ \boldsymbol{R}_{i,lag}]$ is a $(L_{est}\ x\ 3)$ matrix, where the first column, $\boldsymbol{j}$, consists of 1's. The second column is the market's returns $\boldsymbol{R}_m = [R_{m,T_0}, ..., R_{m,T_1}]'$ for the same intervals as the returns of company i. The third column is the lagged returns $\boldsymbol{R}_{i,lag} = [R_{i,T_0-1}, ..., R_{i,T_1-1}]'$. Furthermore, $\boldsymbol{\theta}_i = [\alpha_i\ \beta_i\ \gamma_i]'$ is a $(3\ x\ 1)$ parameter vector and $\boldsymbol{\varepsilon}_i = [\varepsilon_{T_0}, ..., \varepsilon_{T_1}]'$ is a $(L_{est}\ x\ 1)$ residual vector.

We assume the general conditions for ordinary least squares regression hold and that the returns for each company are independently normally distributed. Thus, Equation 2 is a consistent and efficient estimator for the market-model normal returns. Moreover, we find every interval that is 60 minutes before and 90 minutes after the arrival of every news article. These intervals are subtracted from the sample period, and we are left with the estimation window. Further, $\boldsymbol{\theta}_i$ is estimated for each company, which we use to calculate the abnormal returns.

Let $\widehat{\boldsymbol{\varepsilon}}^*_{i,n}$ be the $(L_{event}\ x\ 1)$ vector of estimated abnormal returns for company i and news article n with event window from T_1 to T_2. Using the parameter vector, the vector for abnormal return is given as

$$\widehat{\boldsymbol{\varepsilon}}^*_{i,n} = \boldsymbol{R}^*_i - \widehat{\alpha}_i \boldsymbol{j} - \widehat{\beta}_i \boldsymbol{R}^*_m - \widehat{\gamma}_i \boldsymbol{R}^*_{i,lag} \tag{3}$$

where $\boldsymbol{R}^*_i$ is the return for company i, $\boldsymbol{j}$ is a vector of 1's, $\boldsymbol{R}^*_m$ is the market return in the event window, and $\boldsymbol{R}_{i,lag}$ is the lagged returns. $\widehat{\alpha}_i$, $\widehat{\beta}_i$, and $\widehat{\gamma}_i$ are the estimated parameters from the market model. The abnormal returns are assumed to be jointly normally distributed (Campbell et al., 1996). Moreover, we define $\boldsymbol{\rho}$ as a $(L_{event}\ x\ 1)$ vector with 1's in the positions $\tau_1 - T_1$ to $\tau_2 - T_1$ and zero elsewhere. The cumulative abnormal return (CAR) for news article n from τ_1 to τ_2 is defined as

$$\widehat{CAR}_n(\tau_1, \tau_2) = \boldsymbol{\rho}'\widehat{\boldsymbol{\varepsilon}}^*_{i,n} \tag{4}$$

Furthermore, $\widehat{CAAR}(\tau_1, \tau_2)$ is the average $\widehat{CAR}_n$ across every news article with the same sentiment.[2] Moreover, to test the significance of $\widehat{CAAR}(\tau_1, \tau_2)$, we compute its variance by the following equation

$$Var(\widehat{CAAR}(\tau_1, \tau_2)) = \frac{1}{N^2} \sum_{\tau=\tau_1}^{\tau_2} \sum_{i=1}^{N} \sigma^2_{\varepsilon_i} \tag{5}$$

where $\sigma^2_{\varepsilon_i}$ is the residual standard error from the market-model regression, and N is the total number of news with an equal sentiment. We assume that the variance of $\widehat{CAAR}$ is the same for every interval (Campbell et al., 1996). Hence, we calculate the variance for a single interval and then multiply by the number of intervals we investigate. Furthermore, the null hypothesis of no cumulative abnormal return between τ_1 and τ_2 is tested by using the following test statistics

$$J = \frac{\widehat{CAAR}(\tau_1, \tau_2)}{[Var(\widehat{CAAR}(\tau_1, \tau_2))]^{\frac{1}{2}}} \sim N(0,1) \tag{6}$$

[2] Norwegian Air Shuttle is removed from this part of the analysis due to the extreme return volatility on September 16, 2019, when their bonds were extended. This caused 7% return increase as well as 7% return decrease over two 40-second intervals, which is marked in red in Figure 10 in the Appendix.

3.2 Results

Figure 3 shows the cumulative average abnormal returns around news arrival, split into positive, negative, and neutral sentiment. We find a significant relationship between abnormal returns and news articles labeled as negative. From 60 minutes to 30 minutes before disclosure, there is a significant negative abnormal return of 8 basis points at the 5% level. Then it reverses until the time of disclosure. An insignificant downward movement occurs after disclosure as well, which reverses until about 40 minutes after the news has arrived. Then, abnormal returns decrease significantly until the end of the event window. From the arrival of news to 90 minutes after disclosure, news with negative sentiment has a significant negative abnormal return of 17 basis points at the 1% level. News labeled as positive, show an insignificant downward trend from 60 minutes to 30 minutes before disclosure of 3 basis points. A reversal occurs until the time of news arrival. From news arrival to 90 minutes after, news with positive sentiment has a negative abnormal return of 5 basis points. For news articles labeled with a neutral sentiment, the abnormal return is stable around zero throughout the event window.

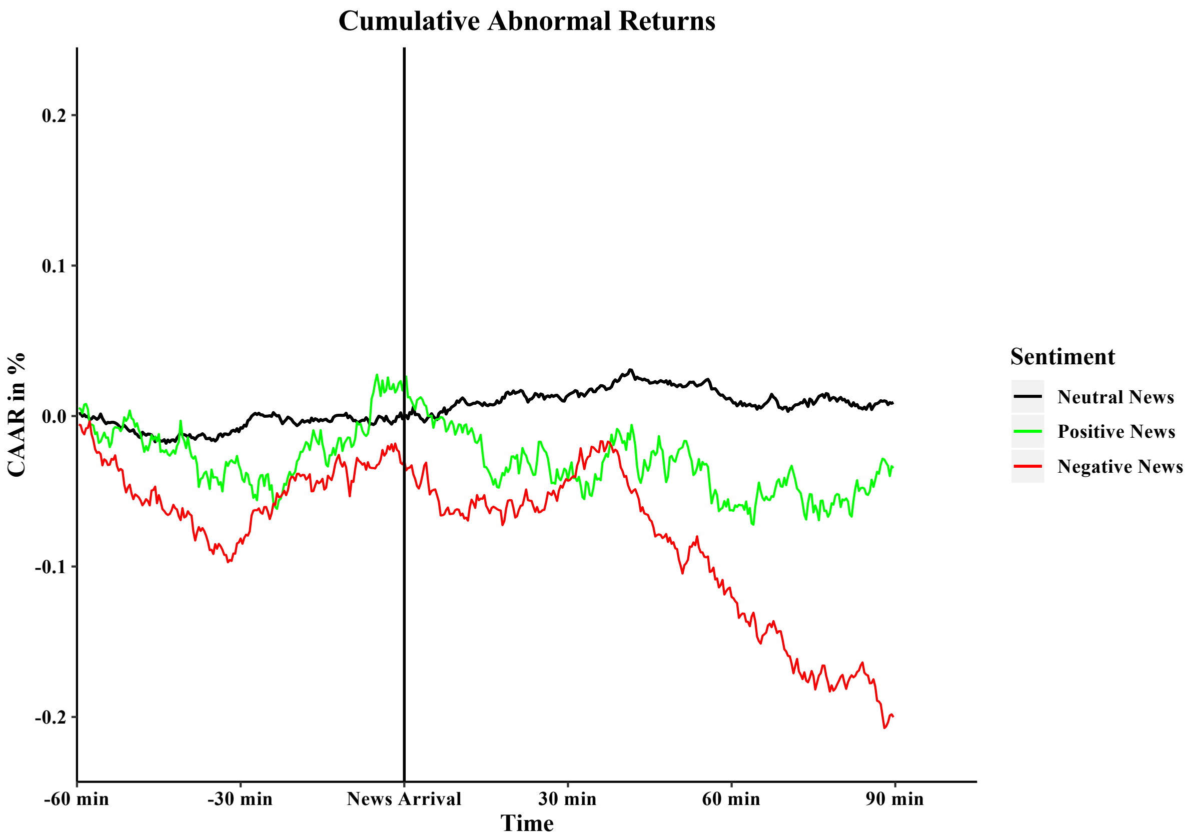

Figure 3: The cumulative average abnormal return calculated for each sentiment.

Table 3: Test statistics for event study.

Interval		Sentiment		
Start (τ_1)	Stop (τ_2)	Negative $J(\tau_1, \tau_2)$	Positive $J(\tau_1, \tau_2)$	Neutral $J(\tau_1, \tau_2)$
-60	-45	-1.80*	-0.46	-1.60
-60	-30	-1.97**	-0.97	-0.56
-60	-15	-0.83	-0.41	-0.18
-60	0	-0.51	0.46	0.01
0	15	-0.70	-1.38	1.10
0	30	-0.24	-1.43	1.06
0	45	-0.67	-1.17	1.52
0	60	-1.41	-1.52	0.71
0	75	-2.11**	-1.18	0.57
0	90	-2.35***	-0.80	0.46

Note: The table shows the test statistics for each sentiment, which is calculated with Equation 6 from τ_1 to τ_2. The intervals are shown in minutes, i.e., $\tau_1 = 0$ and $\tau_2 = 15$ indicates the interval between news disclosure and 15 minutes afterward. *** denotes the significance of the average coefficient estimates at the 1% level, ** at the 5% level, and * at the 10% level.

The test statistic is calculated for several intervals in Table 3. News with negative sentiment has significant negative abnormal returns from 60 minutes to 45 minutes and 30 minutes before news disclosure. There are also significant negative abnormal returns from the news disclosure to 75 minutes and 90 minutes after. News with either positive or neutral sentiment does not have any significant test statistics.

3.3 Discussion

We found that negative news, on average, lead to significant negative abnormal returns for the companies on the OBX Index. This finding supports Tetlock et al. (2008), who find that the fraction of negative words in earnings announcements leads to downward movements in abnormal return on a daily level. While Tetlock et al. (2008) examine scheduled news on a daily aggregation level, we investigate the intraday reactions to unscheduled news. This finding suggests that the relationship between releases of information with negative sentiment and negative abnormal returns hold on a more general level. This finding does, however, contradict Patell and Wolfson (1984), who find that the most significant portion of the price reaction occurs within 15 minutes. One explanation could be that it takes a longer time for the information to be incorporated at the Oslo Stock Exchange, compared to the New York Stock Exchange.

The fact that we see downward market reactions to negative news more than 60 min before the news disclosure is also known in case of periodically scheduled earnings announcements. Kim and Verrecchia (1994) argue that trading before news depends on the degree of information leakages. These results are also supported by Tetlock (2010), who suggests that some investors trade on the information before it becomes public. However, we believe that the trading prior to the news arrival is mainly due to additional information from other news sources. One problem with our analysis is that several sources could publish the same news story at different times. Then it would seem like some investors have access to the information before the rest of the market when they could have received the same information from a news source that published it earlier.

With regard to positive news, the reaction is not significant and has a downward trend after disclosure. This is consistent with De Bondt and Thaler (1985), who find that negative news induces a more significant market reaction than positive news. Moreover, most of the stocks we analyze have been rising in price throughout the sample period, which indicates that the market conditions have been good. This supports Veronesi (1999), who argues that stock prices are likely to overreact to bad news in good times because bad news impacts both expectations about future dividends and the increasing uncertainty. Hence, the fact that we only observe returns reactions to negative news could be due to overreaction.

The results are not consistent Groß-Klußmann and Hautsch (2011) and Smales (2014), who find that abnormal returns are significantly negative (positive) before the disclosure of news with negative (positive) sentiment and insignificant afterward. We argue that there may be multiple reasons for this. The news engine which delivers the news data is not the same. Hence, the linguistic pattern recognition is different, which may affect the labeling process. Thus, the difference in results could be due to different labeling processes for the sentiment, differences in the stock markets, or both. To achieve a better comparison between the Oslo Stock Exchange and the London Stock Exchange or the Australian Securities Exchange, we would have to use the same source for the news data. Moreover, our sample size is much smaller than both Groß-Klußmann and Hautsch (2011) and Smales (2014). This weakens the reliability of our results because there is a bigger chance of getting an unrepresentative outcome.

Lastly, a problem with the methodology is that it implicitly assumes the event is exogenous with respect to the change in the market value of a company. This is particularly relevant in the case of the intraday news flow because a news article could either contain new information or be about an event that has already happened. Another problem with an intraday event study is that returns at short intervals deviate from the normal distribution with higher values of kurtosis and skewness (Aktas, 2008). Moreover, our event study suffers severely from clustering because of overlapping event windows. A news article published in the event window of another article causes a violation of the assumption that the covariance between securities is zero (Campbell et al., 1996). Thus, the normality assumptions we make will not be realistic. With these weaknesses in mind, we are not able to confirm that news with negative sentiment can predict the direction of intraday abnormal returns. Likewise, we are not able to verify that news with positive sentiment does not cause any significant reactions in abnormal returns. Given these points, further research needs to be conducted with a larger sample size and perhaps with a different sentiment indicator to test the first hypothesis further.

4 Volatility and trading activity around news arrivals

4.1 Unconditional properties of volatility and trading activity

In order to get an overview of what to expect from the VAR model, we examine the effects of unscheduled news on volatility and trading activity. We find the unconditional properties of the variables by computing the average reaction of each variable around news arrivals. This means that we do not control for market dynamics and cross-dependencies, such as autocorrelation and cross-correlation between variables. We do this in an attempt to capture important insights about how the disclosure of news affects the intraday market dynamics at the Oslo Stock Exchange. If we assume that all news affects each company equally, we expect to see higher levels, on average, of volatility, money value traded, average trade size, and number of trades around the arrival of news.

We use the same event window as in the event study, with 60 minutes before and 90 minutes after news arrivals with 40-second intervals. Since we have a problem with overlapping event windows, we find this length a reasonable trade-off between capturing pre- and post-effects and reducing the impact of overlapping news articles. For the sake of brevity, we only show the average market reactions across all companies. In order to compare the variables across firms, we standardize all variables as

$$Z_{i,\tau} = \frac{X_{i,\tau} - \mu_i}{\sigma_i} \tag{7}$$

where $Z_{i,\tau}$ is the standardized variable for the company i at interval τ, μ_i is the mean over the sample period, and σ_i is the standard deviation over the same period. Moreover, for each interval, τ, we compute the cross-sectional average market reaction and standard deviation across news using the following formulas

$$\overline{X}_\tau = \frac{1}{N}\sum_{i=1}^{N} X_{i,\tau} \tag{8}$$

$$\sigma_\tau = \sqrt{\frac{1}{(N-1)^2}\sum_{i=1}^{N}(X_{i,\tau} - \overline{X}_\tau)^2} \tag{9}$$

where N is the total number of news for all securities. Equation 9 is used to construct 95% confidence intervals for each time series. Moreover, we use Nadaraya-Watson kernel regression to increase the readability of the output.

The average effects on volatility and money value traded around news disclosure are shown in Figure 4. The left graph shows the average reaction in volatility across every news article. The upward movement starts approximately 60 minutes before disclosure and increases until the news arrives. Afterward, the volatility decreases continuously to a lower level than 60 minutes before disclosure. The right graph shows the average reaction in money value traded across every news article. The upward movement starts 60 minutes before and increases until news disclosure. A new, smaller upward movement occurs after the news has arrived. Afterward, it decreases for about 30 minutes to the same level as it started, before it increases again to 45 minutes after disclosure. The value of money value traded at the end of the event window is close to the level at the start of the event window.

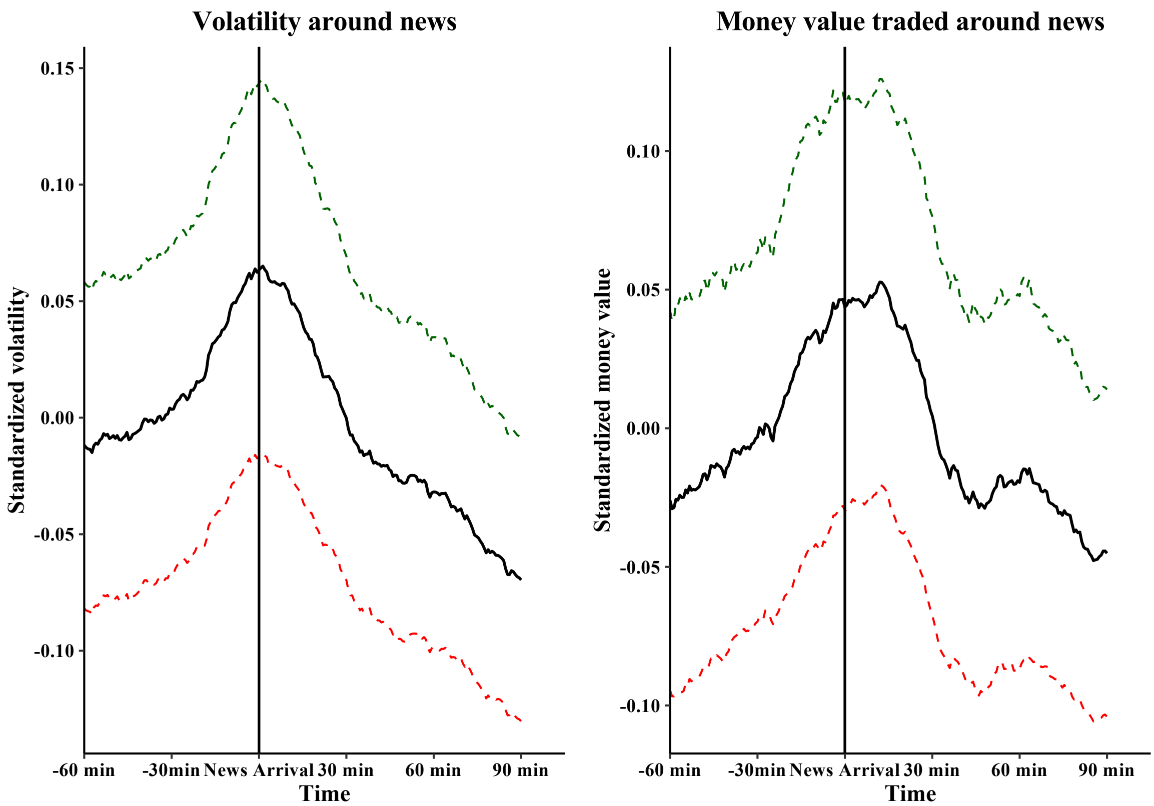

Figure 4: The average reaction of volatility and money value traded around news arrivals. The 95% confidence interval is illustrated with a green dotted line for the upper confidence bound and a red dotted line for the lower confidence bound.

Figure 5 shows the reaction around the news to number of trades and average trade size. The former is shown to the left and has the same characteristics as money value traded. However, the plot for average trade size shows the most irregular reaction around news arrivals. It does, like the other variables, increase from 60 minutes before disclosure until the news arrives, but the reaction is not as centered around the disclosure as the others. Afterward, it decreases for 30 minutes, before it fluctuates up and down for the next 60 minutes. The level of average trade size is higher at the end of the event window compared to the start.

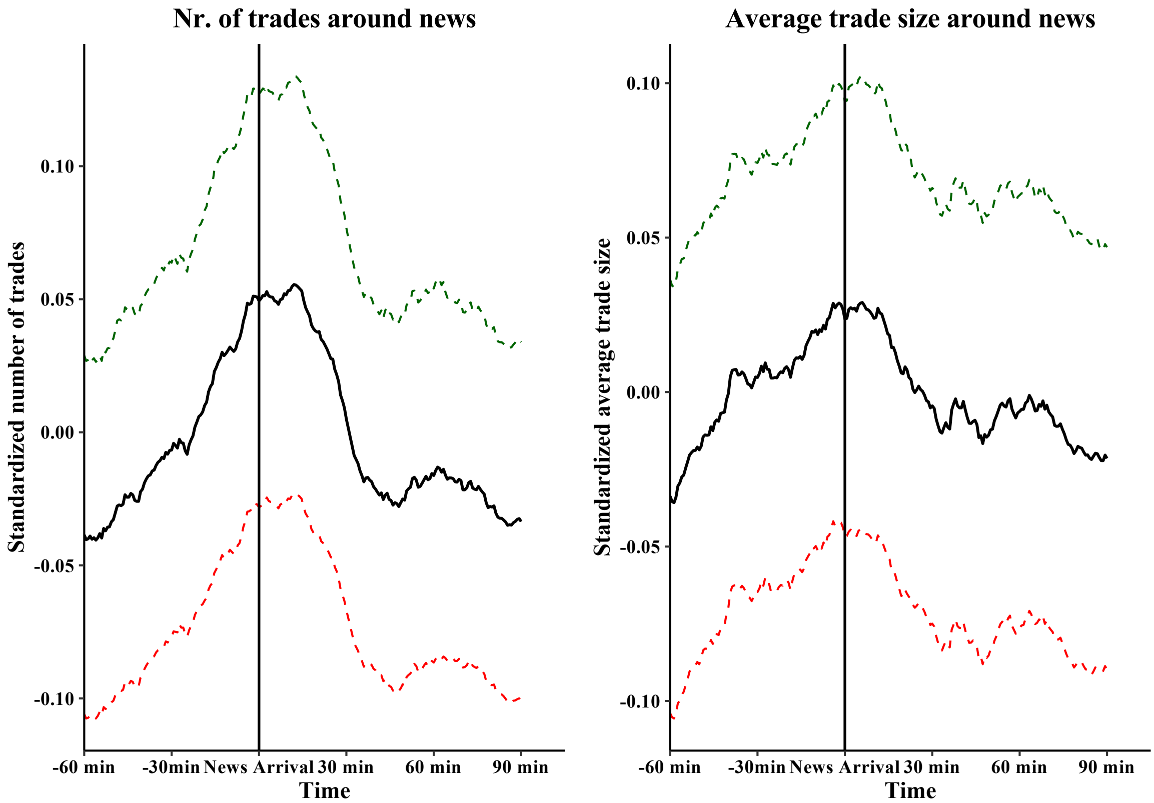

Figure 5: The average reaction in average trade size and number of trades around news arrivals. The 95% confidence interval is illustrated with a green dotted line for the upper confidence bound and a red dotted line for the lower confidence bound.

As observed in the descriptive statistics in Table 2, the companies have a very different number of news articles. To be sure that the average effect is not captured by only a few of the companies, we perform average group-specific means as well. The complete results are found in Section A.1 in the Appendix. The following can be summarized, both volatility and trading activity is qualitatively equal to what we find in Figure 4 and 5. However, because the group-specific means account for the within-group variation, the confidence intervals are more conservative. This is particularly present for average trade size.

4.2 Methodology

The previous section provided initial evidence that unscheduled news leads to market reactions. However, when we check for autocorrelation (Figure 11 in the Appendix), it is clear that the variables show a high degree of persistence. Moreover, we also check for cross-correlation between the variables, which also turns out to be positive (Figure 12 in the Appendix). Based on this, we suspect that some of our findings in Section 4.1 could be spurious and that some of the reactions around news may be spillover effects from other variables. Therefore, we use a VAR model that takes the dependencies and interdependencies between variables into account. A VAR model works by combining multiple univariate autoregressive (AR) models.[3] In short, we create a system of four separate AR models for each of the endogenous variables; volatility, money value traded, average trade size, and number of trades. Compared to a univariate AR model, a VAR model includes lagged variables of every endogenous variable in the system (Lütkepohl, 2005).

First, we check if our time series is stationary, which means that it fluctuate around a constant mean and have a constant variance. We use an augmented Dickey-Fuller test on the four variables for every company (Wooldridge, 2013). The null hypothesis of unit root is discarded at the 1% level for both 40-second and 5-minute intervals, which means the time series is stationary. Therefore, we use a reduced form stationary VAR model, with dummy variables indicating the arrival of news. There are p_1 and p_2 intervals with news dummies before and after the arrival of a news article. We set both p_1 and p_2 to 10, which means that the model captures ten intervals before and after news releases. The following methodology is solely based on Lütkepohl (2005). Our exact specification is defined as follows

$$\boldsymbol{y}_t = \boldsymbol{v} + \sum_{i=1}^{p}(\boldsymbol{A}_i \boldsymbol{y}_{t-p}) + \boldsymbol{C}\boldsymbol{D}_t + \boldsymbol{u}_t \tag{10}$$

where $\boldsymbol{y}_t = [y_{1t}, ..., y_{4t}]'$ is the $(4\ x\ 1)$ output vector, $\boldsymbol{v} = [v_1, ..., v_4]'$ is the $(4\ x\ 1)$ intercept vector, $\boldsymbol{A}_i$ is the $(4\ x\ 4)$ coefficient matrix for lag i, $\boldsymbol{y}_{t-p} = [y_{1t-p}, ..., y_{4t-p}]'$ is the $(4\ x\ 1)$ vector of lagged values of $\boldsymbol{y}_t$ p intervals after t, $\boldsymbol{C}$ is the $(4\ x\ (p_1 + p_2 + 1)) = (4\ x\ 21)$ coefficient matrix for the dummy variables, $\boldsymbol{D}_t = [d_{t-p_1}, ..., d_{t+p_2}]'$ is the $((p_1 + p_2 + 1)\ x\ 1) = (21\ x\ 1)$ vector of

[3] An univariate AR model is a model where the output linearly depends on its own lagged variables (Woodridge, 2013).

dummy variables indicating the arrival of news, and $\boldsymbol{u}_t = [u_{1t}, ..., u_{4t}]'$ is the $(4\ x\ 1)$ vector of white noises. White noise means that $E(\boldsymbol{u}_t) = 0$, $E(\boldsymbol{u}_t \boldsymbol{u}_t') = \sum_u$, and $E(\boldsymbol{u}_t \boldsymbol{u}_s') = 0$ for $s \neq t$. We assume the covariance matrix $\sum_u$ is nonsingular, which means that $det(\sum_u) \neq 0$. This assumption is necessary to be able to solve the system of equations (Lütkepohl, 2005).

Furthermore, we use the Akaike's Information Criterion (AIC) to determine the optimal number of lags p. For a $VAR(p)$ model, the AIC is defined as

$$AIC(p) = ln[det(\sum_u^{\sim}(m))] + \frac{2pK^2}{T} \tag{11}$$

where $\sum_u^{\sim}$ is the estimated covariance matrix, p is the number of lagged endogenous variables, and pK^2 is the number of freely estimated parameters. The optimal number of lags, $\widehat{p}(AIC)$, choose the number of lags so that $AIC(p)$ is minimized (Lütkepohl, 2005). Thus, the closer the determinant of the covariance matrix is to zero, the more the criterion is minimized. Also, as the number of lags increases, the greater the criterion becomes. In sum, the criterion is a trade-off between how well the parameters are estimated and how many lags are included.

To test how the endogenous variables respond to the arrival of news, we perform an impulse response analysis. It works by tracing out the effect of an exogenous shock to the system. To isolate the effect of news arrivals, we assume the endogenous variables are at their mean value before the impulse at time $t = 0$, which is zero since the variables have been standardized (Lütkepohl, 2005). Moreover, one of the news dummies is set equal to one, i.e., $\boldsymbol{D}_t = [0, ..., 1, ..., 0]'$. This simulates a news arrival in the interval corresponding to the dummy that is activated.

We use Equation 10 to compute $\boldsymbol{y}_0, ..., \boldsymbol{y}_{10}$. Thus, we trace out how the endogenous variables respond ten intervals after the initial shock. Since the endogenous variables are zero before $t = 0$, the output from the system at $t = 0$ is only based on the "news shock". Thus, $\boldsymbol{y}_0 = \boldsymbol{C}\boldsymbol{D}_t$. After the initial impulse, the lagged values affect the next period, and we have $\boldsymbol{y}_1 = \boldsymbol{A}_1 \boldsymbol{y}_0$. At $t = 2$ we compute $\boldsymbol{y}_2 = \boldsymbol{A}_1 \boldsymbol{y}_1 + \boldsymbol{A}_2 \boldsymbol{y}_0$, and so forth until $t = 10$.

4.3 Results

In this section, we first report the results from the VAR models based on 40-second intervals. Next, we report the results from the VAR models based on 5-minute intervals. Every VAR model has the same four variables from Section 4.1 as its endogenous variables; volatility, money value traded, average trade size, and number of trades.

Table 4 shows the average 40-second VAR results for the standardized variables across all companies. Each column represents an $AR(p)$ model, which is regressed on its own lagged variables, every other endogenous lagged variable, and dummies for news arrival. Each dummy variable represents a 40-second interval relative to the arrival of the news item. For the sake of brevity, we only report two lags for the endogenous variables, and three lag and lead dummy variables for news arrival.

First, the $AR(p)$ model for money value traded is significantly dependent on its own lagged variables, lagged variables of volatility, and lagged variables of number of trades. Second, the $AR(p)$ model for volatility is significantly dependent on its own lagged variables, the second-order lag of money value traded, and first-order lag of average trade size and number of trades. Third, the $AR(p)$ model for average trade size is significantly dependent on its own lagged variables, lagged variables of both volatility and number of trades, and the first-order lag of money value traded. Lastly, the $AR(p)$ model for number of trades is significantly dependent on its own lagged variables, lagged variables of volatility, and the first-order lag of money value traded. None of the news dummies are significant for any of the four $AR(p)$ models, which means that news does not, on average, significantly increase any of the endogenous variables at any of the ten 40-second intervals before and after the arrivals of news.

Although we cannot capture significant market reactions on average for all companies, we do find significant market reactions for almost all of the companies but at different intervals. Figure 6 shows the total number of significant news dummies for each interval across every company. The effects of each variable are shown over ten 40-second intervals before and after the news disclosure. We observe that the significant news dummies are not centered around the news disclosure. The significant dummies are well distributed over the event window, for both volatility, money value traded, and number of trades. On the contrary, there are few significant dummies around the closest intervals of news arrival for average trade size.

Table 4: Average VAR results with 40-second intervals.

Dynamics		Money value	Volatility	Average trade	Nr. of trades
Constant	c	-0.001	-0.001	0.000	-0.001
		(0.005)	(0.005)	(0.005)	(0.004)
Money value	mv_{t-1}	0.144***	-0.023	0.027**	0.045***
		(0.010)	(0.010)	(0.011)	(0.010)
	mv_{t-2}	0.060***	0.025**	0.004	0.010
		(0.010)	(0.010)	(0.011)	(0.010)
Volatility	$vola_{t-1}$	0.062***	0.197***	0.030***	0.041***
		(0.005)	(0.005)	(0.006)	(0.005)
	$vola_{t-2}$	0.014**	0.047***	0.014**	0.013**
		(0.005)	(0.005)	(0.006)	(0.005)
Average Trade	ats_{t-1}	-0.029	0.080***	0.057***	-0.026
		(0.006)	(0.006)	(0.006)	(0.006)
	ats_{t-2}	-0.013	-0.029	0.038***	-0.011
		(0.006)	(0.006)	(0.006)	(0.006)
Nr. of trades	n_{t-1}	0.019*	0.156***	0.031***	0.126***
		(0.010)	(0.010)	(0.010)	(0.009)
	n_{t-2}	0.018**	-0.033	0.018*	0.079***
		(0.010)	(0.010)	(0.010)	(0.009)
Dummy leads	d_{t+3}	0.010	-0.012	-0.011	0.036
		(0.239)	(0.236)	(0.251)	(0.234)
	d_{t+2}	0.035	0.075	0.110	0.107
		(0.239)	(0.236)	(0.251)	(0.234)
	d_{t+1}	-0.093	0.018	-0.095	-0.042
		(0.239)	(0.236)	(0.251)	(0.234)
News arrival	d_t	0.132	0.035	0.001	0.043
		(0.239)	(0.236)	(0.251)	(0.234)
Dummy lags	d_{t-1}	-0.037	0.003	0.092	-0.061
		(0.239)	(0.236)	(0.251)	(0.234)
	d_{t-2}	0.119	-0.019	-0.010	0.098
		(0.239)	(0.236)	(0.251)	(0.234)
	d_{t-3}	0.001	0.120	-0.046	0.059
		(0.239)	(0.236)	(0.251)	(0.234)

Note: The table provides the average estimation results for the VAR models, as outlined in Equation 10. Estimates are provided for the dynamics of the endogenous variables, together with the exogenous news dummies. Reported coefficients are the average of the coefficients for each individual company with average standard errors given in the parentheses below. Significance is reported based on the average of the t-values. *** denotes the significance of the average coefficient estimates at the 1% level, ** at the 5% level, and * at the 10% level.

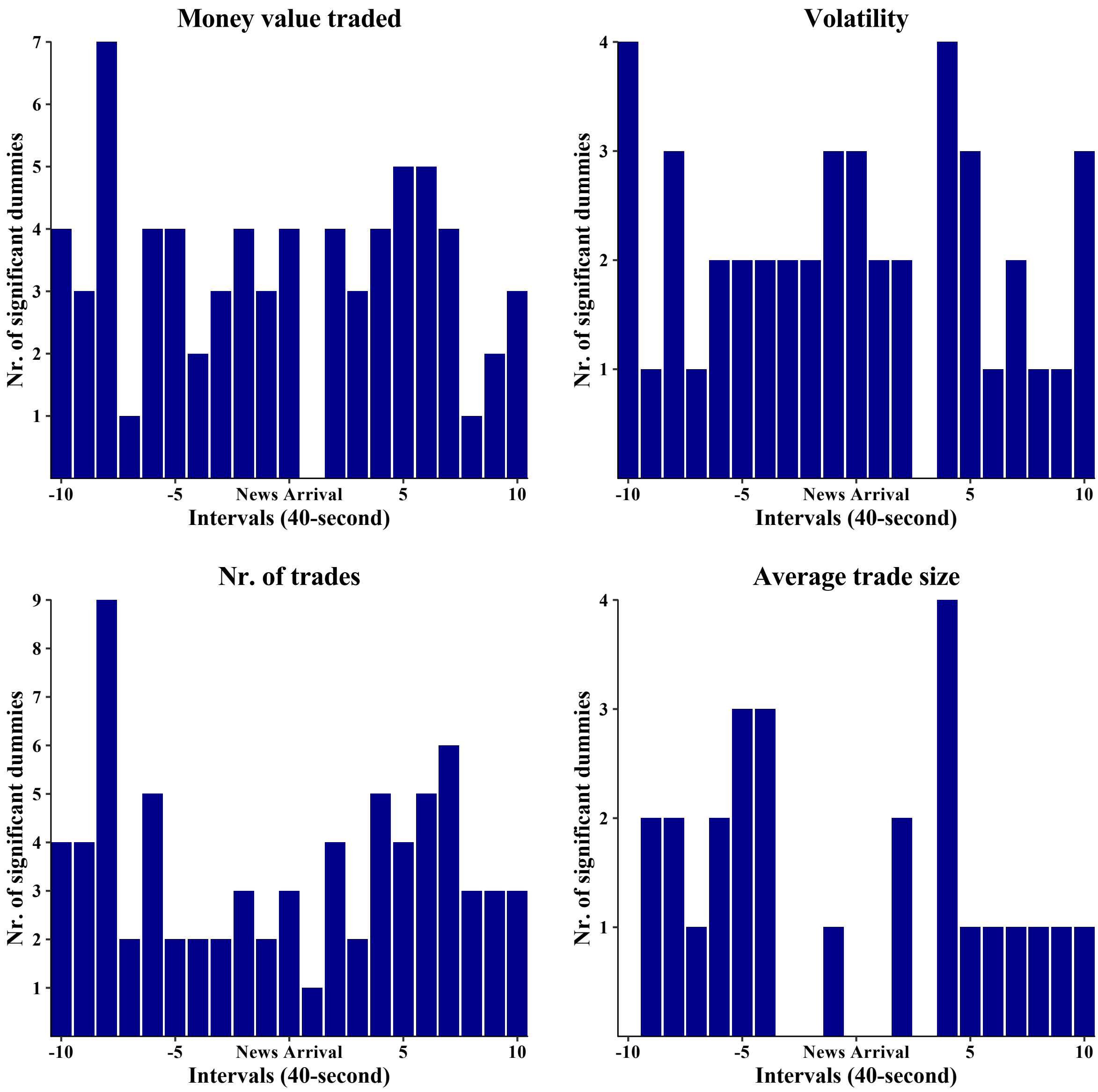

Figure 6: The total number of significant news dummies across every company from the VAR models with 40-second intervals. The reported dummies are at least significant at the 5% level.

Nevertheless, as shown in Table 4, we are not able to detect significant average reaction for news arrivals. However, Section 4.1 did indicate a reaction starting before the news disclosure. Moreover, the companies, as shown in Table 2, differ in number of trades. As discussed, Hautsch et al. (2013) address problems with time series on high-frequency data because of an excess amount of zero observations. To address the excess amount of zero observations and short intervals of exogenous variables, we perform a VAR model where we aggregate the intervals to five minutes. Thus, we capture 50 minutes before and after the arrival of news.

Table 5: Average VAR results with 5-minute intervals.

Dynamics		Money value	Volatility	Average trade	Nr. of trades
Constant	c	-0.006	-0.004	-0.002	-0.005
		(0.011)	(0.013)	(0.013)	(0.011)
Money value	mv_{t-1}	0.216***	0.129***	-0.058	-0.021
		(0.035)	(0.040)	(0.041)	(0.033)
	mv_{t-2}	0.126***	0.096**	0.027	0.081**
		(0.036)	(0.041)	(0.042)	(0.034)
Volatility	$vola_{t-1}$	0.085***	0.109***	0.025	0.085***
		(0.013)	(0.015)	(0.015)	(0.013)
	$vola_{t-2}$	0.002	0.032**	0.011	0.000
		(0.013)	(0.015)	(0.016)	(0.013)
Average Trade	ats_{t-1}	-0.017	-0.021	0.181***	0.011
		(0.015)	(0.016)	(0.017)	(0.014)
	ats_{t-2}	-0.041	-0.079	0.073***	-0.046
		(0.015)	(0.017)	(0.017)	(0.014)
Nr. of trades	n_{t-1}	0.113***	0.012	0.106***	0.374***
		(0.032)	(0.037)	(0.038)	(0.031)
	n_{t-2}	-0.032	-0.122	-0.063	0.025
		(0.034)	(0.039)	(0.04)	(0.032)
Dummy leads	d_{t+3}	-0.012	-0.007	-0.024	0.022
		(0.211)	(0.237)	(0.247)	(0.203)
	d_{t+2}	0.012	-0.165	0.091	-0.021
		(0.211)	(0.237)	(0.247)	(0.203)
	d_{t+1}	0.118	0.069	-0.001	0.067
		(0.211)	(0.237)	(0.247)	(0.203)
News arrival	d_t	0.063	-0.003	0.069	0.074
		(0.211)	(0.237)	(0.247)	(0.203)
Dummy lags	d_{t-1}	0.484*	0.077	0.177	0.350
		(0.210)	(0.236)	(0.247)	(0.202)
	d_{t-2}	0.469*	0.474*	0.174	0.326
		(0.210)	(0.236)	(0.246)	(0.202)
	d_{t-3}	0.224	0.449	0.076	0.120
		(0.211)	(0.237)	(0.247)	(0.203)

Note: The table provides the average estimation results for the VAR models, as outlined in Equation 10. Estimates are provided for the dynamics of the endogenous variables, together with the exogenous news dummies. Reported coefficients are the average of the coefficients for each individual company with average standard errors given in the parentheses below. Significance is reported based on the average of the t-values. *** denotes the significance of the average coefficient estimates at the 1% level, ** at the 5% level, and * at the 10% level.

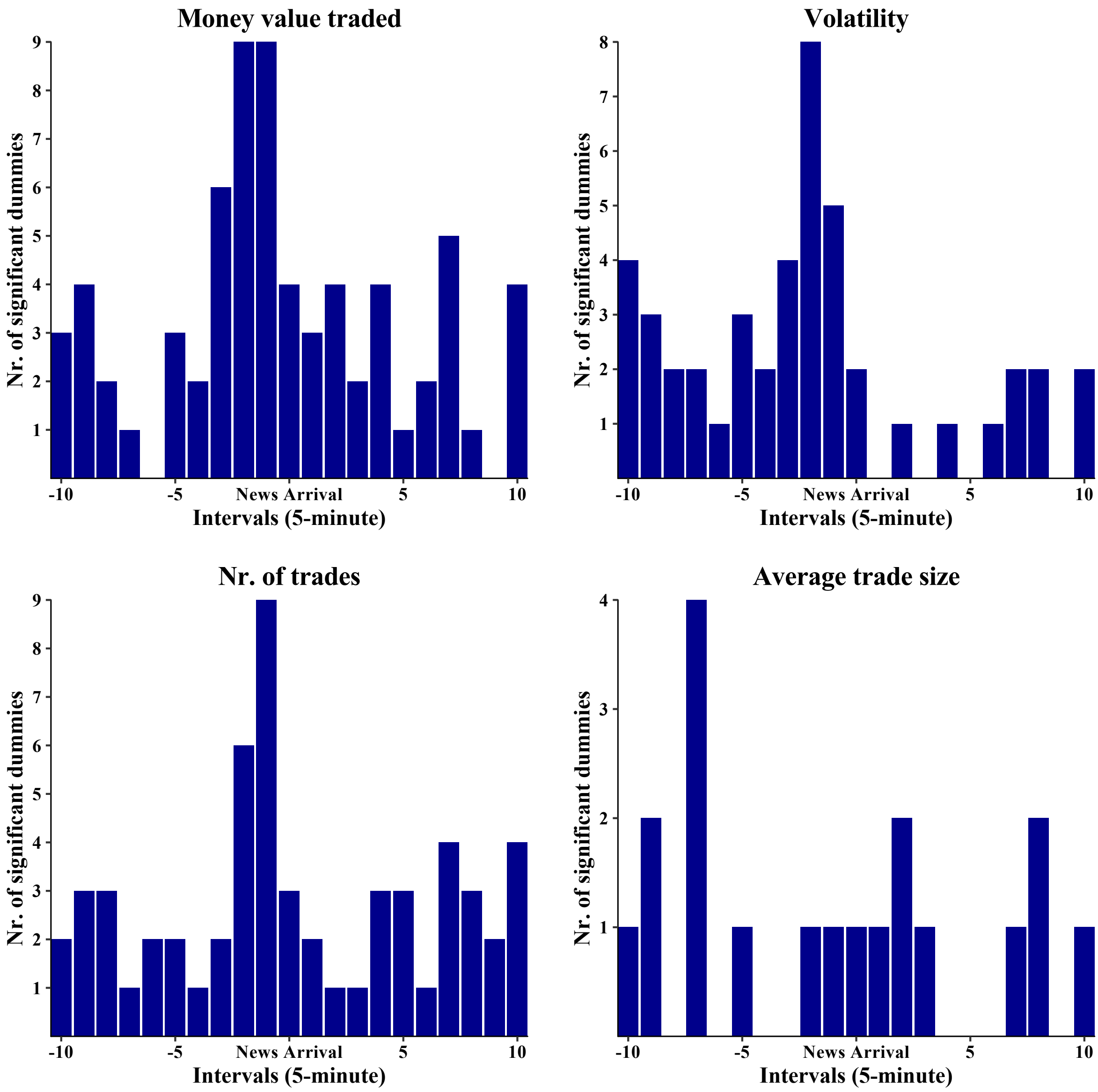

Figure 7: The total number of significant news dummies across every company from the VAR models with 5-minute intervals. The reported dummies are at least significant at the 5% level.

Table 5 shows the average VAR results for the standardized variables across all companies. First, we observe that the news dummies d_{t-1} and d_{t-2} are both significant at the 10% level for money value traded. The causal interpretation of the standardized averaged money value traded is as follows: on average, money value traded increases by 0.47 standard deviations ten minutes prior to the news disclosure. Further, money value traded increases on average by 0.48 standard deviations five minutes prior to the news disclosure. Second, the news dummy d_{t-2} is significant at the 10% level for volatility. That is, on average, volatility increase by 0.47 standard deviations ten minutes prior to the news disclosure. Third, every endogenous variable is still significantly dependent on its own lagged variable. Lastly, with regard to the cross-dependencies, some of the significant relationships in the 40-second VAR model are no longer significant with 5-minute intervals. Volatility is no longer dependent on the lagged variables of average trade size and number of trades. Average trade size is no longer dependent on the lagged variables of money value traded and volatility.

Figure 7 shows number of significant news dummies in the event window for 5-minute intervals for every variable. First, compared to Figure 6, the news dummies are centered around news arrival. We observe that money value traded has the most significant news dummies five and ten minutes prior to the news arrival, where both intervals have nine significant news dummies. Likewise, volatility has the most significant news dummies ten minutes prior to the news arrival, with eight significant news dummies. These results are consistent with the results from Table 5, where both money value traded and volatility have significant news dummies in the intervals with the largest number of significant news dummies across companies. Furthermore, number of trades has at most nine significant news dummies five minutes prior to the news arrival. Average trade size has at most four significant news dummies seven intervals before the news arrival. Additionally, average trade size has the fewest number of significant news dummies of all the endogenous variables.

Both Figure 6 and Figure 7 show a tendency of significant reactions in news dummies on money value traded, volatility, average trade size and number of news. To be sure that there are not only a few companies that capture these effects, we also inspect the individual VAR models with 5-minute intervals to find out which companies have the most significant news dummies during the event window. Table 6 shows that there are differences between companies regarding how many significant news dummies there are during the event window. The companies are sorted in descending order by the total number of significant news dummies. SUBC has the most, with a total of 34 significant dummies. The company with the least is DNO, which has zero significant dummies. Moreover, the ten companies with the most significant news dummies account for 70% of the total number of significant news dummies.

Table 6: Number of significant news dummies in 5-minute VAR model.

Ticker	Volatility	Money value	Average trade	Nr. of trades	Sum
SUBC	8	11	5	10	34
STB	7	9	1	7	24
TOM	10	5	0	7	22
SALM	6	7	1	5	19
TEL	8	7	0	4	19
GJF	3	6	2	4	15
NEL	1	4	3	7	15
LSG	2	5	2	5	14
DNB	4	5	3	1	13
NHY	1	5	2	4	12
PGS	1	4	0	6	11
AKER	1	4	1	4	10
BWO	1	4	2	3	10
NAS	2	4	0	3	9
ORK	3	3	1	2	9
YAR	1	3	0	3	7
EQNR	1	3	1	1	6
SCHA	2	2	0	1	5
BAKKA	1	0	3	0	4
AKERBP	1	1	0	1	3
GOGL	0	2	1	0	3
TGS	1	0	1	1	3
ELK	1	0	0	1	2
MOWI	0	0	0	1	1
DNO	0	0	0	0	0
Sum	66	94	29	81	270

Note: The table shows how many significant news dummies each company has at the 5% level with the 5-minute VAR specification. The dummies capture $\pm$ 50 minutes relative to the news disclosure. The first column shows the companies sorted in descending order by the total number of signifiant nummies. The next four columns show the dependent variables, and the fifth column shows the total number of significant news dummies across all dependent variables.

To provide further insights into news induced market responses in a dynamic system, we use an impulse response analysis. We define a "news shock" as a shock to the news dummy d_{t-2}. This means that the reaction in the endogenous variables is simulated as if the news article came ten minutes earlier. Figure 8 shows the impulse response to news-induced dummy variable changes based on the averaged VAR estimates in Table 5. Similar to the average reaction in the endogenous variables in Section 4.1, we observe that volatility has a sharp spike, which subsides quickly. In contrast, money value traded and number of trades have a lower response, but the relative effect lasts longer. Lastly, average trade size has a small reaction that subsides quickly.

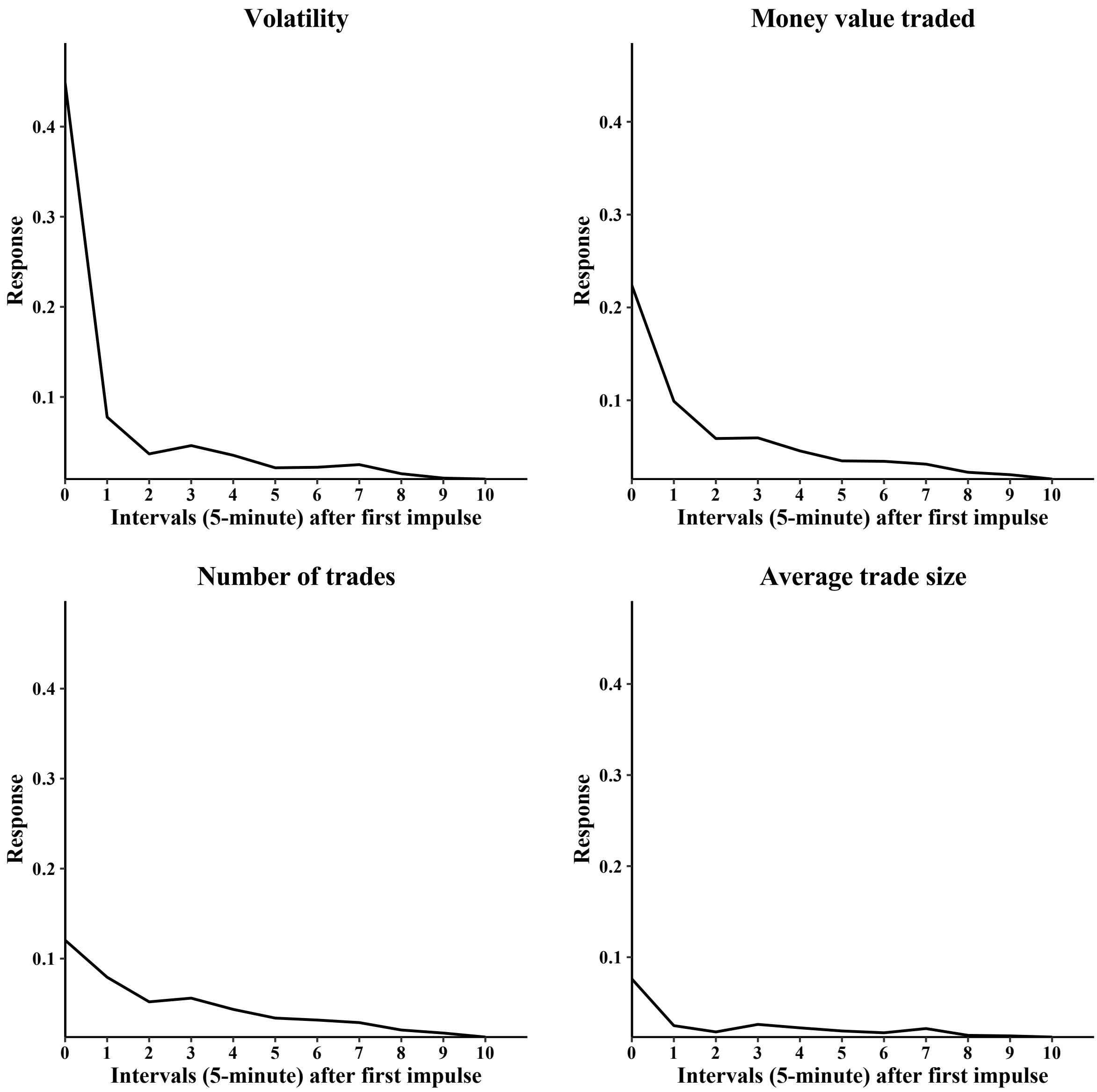

Figure 8: Impulse response functions from a shock in the news dummy d_{t-2} on endogenous variables. The impulse response functions are based on the average VAR estimates from the model with 5-minute intervals. Hence, the impulse response shows how the endogenous variables react 50 minutes after the initial impulse.

4.4 Robustness

In order to validate our results, we perform several robustness tests. The complete set of tests can be found in (Section A.1 in the Appendix).

A common characteristic of intraday data series is the existence of a distinct U-shaped pattern in volatility and trading activity (Andersen & Bollerslev, 1997). Although we discarded 15 minutes at the start and end of the trading day to get rid of opening and close effects, we still observe intraday seasonality (see Figure 13 and 14 in the Appendix). To account for this, we create dummy variables indicating the interval of the trading day, with the base dummy set to the first interval in the trading day. Then, we re-run the VAR model with 5-minute intervals, including the interval dummies. The results are reported in Table 7. We observe that the same relationship between news arrival and volatility and money value traded are present. However, d_{t-2} is significant at the 5% level on volatility compared to 10% in Table 5. Moreover, the coefficients of d_{t-3} and d_{t-2} on money value traded are slightly reduced, but still significant at the 10% level. Moreover, we observe that the constant term is higher for all endogenous variables compared to Table 5. This is because the base dummy captures the opening effect, as illustrated in Table 7.

Furthermore, empirical findings in finance show that there exists seasonality among the days of the week (Doyle & Chen, 2009). To account for this, we include dummy variables indicating the day of the week. The results do not change; however, the second lag of news arrival, d_{t-1}, is now significant at the 5% level on volatility compared to 10% in Table 9. Finally, to ensure that we capture effects induced by firm-specific news, we randomly assign news articles to each company and re-run the VAR models with 5-minute intervals. As the results show in Table 8, we cannot find any significant relationship on average between news arrival and any of the endogenous variables.

Table 7: Average VAR results with 5-minute intervals adjusted for intraday seasonality.

Dynamics		Money value	Volatility	Average trade	Nr. of trades
Constant	c	0.272***	1.529***	-0.435	0.074
		(0.094)	(0.107)	(0.108)	(0.087)
Money value	mv_{t-1}	0.259***	0.060	0.098**	0.054*
		(0.035)	(0.041)	(0.041)	(0.033)
	mv_{t-2}	0.054*	-0.054	0.000	0.012
		(0.036)	(0.042)	(0.042)	(0.034)
Volatility	$vola_{t-1}$	0.071***	0.087***	0.004	0.064***
		(0.014)	(0.016)	(0.016)	(0.013)
	$vola_{t-2}$	0.009	0.015	-0.013	0.011
		(0.014)	(0.016)	(0.016)	(0.013)
Average Trade	ats_{t-1}	-0.043	0.024	0.083***	-0.034
		(0.016)	(0.018)	(0.018)	(0.015)
	ats_{t-2}	-0.003	0.017	0.071***	-0.011
		(0.016)	(0.018)	(0.018)	(0.015)
Nr. of trades	n_{t-1}	0.048	0.085**	-0.057	0.271***
		(0.034)	(0.039)	(0.040)	(0.032)
	n_{t-2}	0.060	0.064	0.019	0.114***
		(0.035)	(0.041)	(0.041)	(0.033)
Dummy leads	d_{t+3}	-0.016	0.022	-0.043	0.015
		(0.205)	(0.232)	(0.237)	(0.191)
	d_{t+2}	0.005	-0.136	0.055	-0.026
		(0.205)	(0.232)	(0.237)	(0.191)
	d_{t+1}	0.130	0.091	-0.025	0.086
		(0.205)	(0.232)	(0.237)	(0.191)
News arrival	d_t	0.077	0.026	0.054	0.085
		(0.205)	(0.232)	(0.237)	(0.191)
Dummy lags	d_{t-1}	0.481*	0.092	0.151	0.351
		(0.204)	(0.232)	(0.237)	(0.191)
	d_{t-2}	0.451*	0.475**	0.127	0.314
		(0.204)	(0.231)	(0.236)	(0.190)
	d_{t-3}	0.229	0.431	0.078	0.131
		(0.205)	(0.232)	(0.237)	(0.191)
Intraday seasonality		Yes	Yes	Yes	Yes

Note: The table provides the average estimation results for the VAR models, with dummy variables to control for intraday seasonality. Estimates are provided for the dynamics of the endogenous variables, together with the exogenous news dummies. Reported coefficients are the average of the coefficients for each individual company with average standard errors given in the parentheses below. Significance is reported based on the average of the t-values. *** denotes the significance of the average coefficient estimates at the 1% level, ** at the 5% level, and * at the 10% level.

4.5 Discussion

We find that both volatility and money value traded increase significantly in-front of news arrival, consistent with Groß-Klußmann and Hautsch (2011) and Smales (2014). This supports the mixture of distribution hypothesis, to the extent that price change and trading volume change are driven by the same underlying information arrival process (Clark, 1973). Thus, our results are in line with what Kalev et al. (2004) find. However, while they find that the total number of news articles significantly affect volatility, our results suggest that the increase may come from each news article, and not the arrival rate alone. Moreover, our results support that investors display attention-grabbing behavior (Barber & Odean, 2008). In sum, these results support (Fang & Peress, 2009), in the way that media can mitigate asymmetrical information between investors and affect stock prices even when the news is stale.

In contrast to Groß-Klußmann and Hautsch (2011) and Smales (2014), neither volatility nor money value traded is significant at the news arrival, but the results suggest that the significant reaction happens five and ten minutes before. We argue that there may be multiple reasons for this. First, there could be larger information leakages (Kim and Verrecchia (1994) and Tetlock (2010)) at the Oslo Stock Exchange, compared to the London Stock Exchange and the Australian Securities Exchange. Second, our news articles could contain information about news published earlier. In other words, the significant reaction we capture could come from investors reacting to news articles out of our sample. Furthermore, we cannot rule out the fact that a reverse causality problem may be present. In other words, the news articles could contain information about, for example, a jump in volatility or increased trading activity. Thus, the news write about market movements which imply that the news variables are not exogenous.[4]

Moreover, the fact that we find insignificant news reactions on 40-second intervals may be because there are differences in what time the different stocks incorporate information. Figure 6 shows that the significant news dummies are well distributed over the short event window. Moreover, Table 6 showed that there are large differences in how the different stocks respond to the news. Hence, the results suggest that there are news induced reactions on volatility and money value traded at 40-second intervals, but that the effect is diffuse across the cross-section of stocks. There may be several explanations for this. First, there could be differences between the companies in the OBX Index that we have not adjusted for. For example, Brennan, Jegadeesh,

[4]Violates the zero conditional mean assumption in OLS. Thus, OLS is no longer unbiased and consistent.

and Swaminathan (1993) argue that companies with less informed investors use longer time to incorporate information. Moreover, Shleifer and Vishny (1997) argue that reactions in low-risk firms appear before high-risk firms because arbitrageurs avoid firms with high idiosyncratic risk. Consequently, idiosyncratic differences between the companies can be the reason why the effects are not significant. Second, we treat all news articles as equal. Therefore, we cannot rule out the fact that some companies are affected by a higher number of market-moving news.

To conclude, the result supports our hypothesis that intraday volatility and trading activity increase around the arrival of news stories from the financial media.

5 Conclusion

Processing and analyzing the overall news feed for a specific company is challenging due to the continuous flow of information from mass media. It is not trivial to filter out relevant information. Hence, it has been challenging to find a significant link between high-frequency market dynamics and the intraday news flow. This is why the majority of previous studies focus on easily identifiable and homogenous news.

By using pre-processed news data based on linguistic pattern recognition combined with high-frequency stock data from the Oslo Stock Exchange, we tested how the intraday news flow affects high-frequency market movements. First, we used an event study on stock data aggregated into 40-second intervals. The results suggested that negative news articles led to significant negative cumulative abnormal returns. Moreover, the average effects of volatility and trading activity around news arrivals suggested a reaction in both variables when a news article is published. Then, to take the strong dependencies and interdependencies into account, we used a high-frequency VAR model. With 40-second intervals, we did not find a significant reaction to news arrivals. However, by using 5-minute intervals, we avoided the problem of excess zero observations in the data. In this case, we found a significant positive reaction in both volatility and money value traded ten minutes prior to the news arrival. In addition, money value traded reacted significantly positive five minutes before the news arrival.

Our results suggest that negative news induces more significant price reactions than positive news. Moreover, our findings propose that unscheduled news affects both intraday volatility and trading activity for the companies at the Oslo Stock Exchange. The results from the VAR model are robust when we control for weekday effects and intraday seasonality pattern, which supports our findings. However, there are limitations to our approaches. First, our sentiment indicator may not be optimal for capturing the tone of financial news. Thus, this can be an explanation of why we cannot correctly identify the direction of cumulative abnormal returns. Second, the fact that volatility and trading activity reacts five and ten minutes before a news arrival could be because of a reverse causality problem being present. We suggest further analysis of a larger sample. In addition, looking at the market reaction induced by unscheduled news between different stocks would be of interest, as there seem to be idiosyncratic reactions through the cross-section of stocks at the Oslo Stock Exchange.

References

Aktas, E. (2008). Intraday stock returns and performance of a simple market model. *Applied Financial Economics*, *18*(18), 1475–1480.

Andersen, T. G., & Bollerslev, T. (1997). Intraday periodicity and volatility persistence in financial markets. *Journal of Empirical Finance*, *4*(2-3), 115–158.

Antweiler, W., & Frank, M. Z. (2004). Is all that talk just noise? the information content of internet stock message boards. *The Journal of Finance*, *59*(3), 1259–1294.

Bachmann, K., Enrico, D. G. G., & Hens, T. (2018). *Behavioral finance for private banking: From the art of advice to the science of advice* (2nd ed.). Wiley.

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

Barndorff-Nielsen, O. E., & Shephard, N. (2002). Econometric analysis of realized volatility and its use in estimating stochastic volatility models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *64*(2), 253–280.

Berry, T. D., & Howe, K. M. (1994). Public information arrival. *The Journal of Finance*, *49*(4), 1331–1346.

Boudoukh, J., Feldman, R., Kogan, S., & Richardson, M. (2013). Which news moves stock prices? a textual analysis. *NBER Working Paper*, *18725*.

Brennan, M. J., Jegadeesh, N., & Swaminathan, B. (1993). Investment analysis and the adjustment of stock prices to common information. *Review of Financial Studies*, *6*(4), 799–824.

Campbell, J. Y., Lo, A. W., & MacKinlay, A. C. (1996). *The econometrics of financial markets* (2nd ed.). Princeton University Press.

Chen, H., De, P., Hu, Y. (, & Hwang, B.-H. (2014). Wisdom of crowds: The value of stock opinions transmitted through social media. *Review of Financial Studies*, *27*(5), 1367–1403.

Clark, P. K. (1973). A subordinated stochastic process model with finite variance for speculative prices. *Econometrica*, *41*(1), 135–155.

Cont, R. (2001). Empirical properties of asset returns: Stylized facts and statistical issues. *Quantitative Finance*, *1*(2), 223–236.

Cutler, D. M., Poterba, J. M., & Summers, L. H. (1989). What moves stock prices? *The Journal of Portfolio Management*, *15*(3), 4–12.

De Bondt, W. F. M., & Thaler, R. H. (1985). Does the stock market overreact? *The Journal of Finance*, *40*(3), 793–805.

Dougal, C., Engelberg, J., García, D., & Parsons, C. A. (2012). Journalists and the stock market. *Review of Financial Studies*, *25*(3), 639–679.

Doyle, J. R., & Chen, C. H. (2009). The wandering weekday effect in major stock markets. *Journal of Banking & Finance*, *33*(8), 1388–1399.

Duan, J.-C., Härdle, W. K., & Gentle, J. E. (2012). *Handbook of computational finance*. Springer.

Engelberg, J. E., & Parsons, C. A. (2011). The causal impact of media in financial markets. *The Journal of Finance*, *66*(1), 67–97.

Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *The Journal of Finance*, *25*(2), 383–417.

Fang, L., & Peress, J. (2009). Media coverage and the cross-section of stock returns. *The Journal of Finance*, *64*(5), 2023–2052.

García, D. (2013). Sentiment during recessions. *The Journal of Finance*, *68*(3), 1267–1300.

Groß-Klußmann, A., & Hautsch, N. (2011). When machines read the news: Using automated text analytics to quantify high frequency news-implied market reactions. *Journal of Empirical Finance*, *18*(2), 321–340.

Harris, L. (1987). Transaction data tests of the mixture of distributions hypothesis. *The Journal of Financial and Quantitative Analysis*, *22*(2), 127–141.

Harris, M., & Raviv, A. (1993). Differences of opinion make a horse race. *Review of Financial Studies*, *6*(3), 473–506.

Hautsch, N., Malec, P., & Schienle, M. (2013). Capturing the zero: A new class of zero-augmented distributions and multiplicative error processes. *Journal of Financial Econometrics*, *12*(1), 89–121.

Heston, S. L., & Sinha, N. R. (2017). News vs. sentiment: Predicting stock returns from news stories. *Financial Analysts Journal*, *73*(3), 67–83.

Huberman, G., & Regev, T. (2001). Contagious speculation and a cure for cancer: A nonevent that made stock prices soar. *The Journal of Finance*, *56*(1), 387–396.

Kalev, P. S., Liu, W.-M., Pham, P. K., & Jarnecic, E. (2004). Public information arrival and volatility of intraday stock returns. *Journal of Banking & Finance*, *28*(6), 1441–1467.

Kim, O., & Verrecchia, R. E. (1994). Market liquidity and volume around earnings announcements. *Journal of Accounting and Economics*, *17*(1-2), 41–67.

Larsen, V. H., & Thorsrud, L. A. (2017). Asset returns, news topics, and media effects. *Norges Bank Research*, *17*.

Lee, C. M. (1992). Earnings news and small traders. *Journal of Accounting and Economics*, *15*(2-3), 265–302.

Lütkepohl, H. (2005). *New introduction to multiple time series analysis*. Springer.

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

Odean, T. (1999). Do investors trade too much? *American Economic Review*, *89*(5), 1279–1298.

Patell, J. M., & Wolfson, M. A. (1984). The intraday speed of adjustment of stock prices to earnings and dividend announcements. *Journal of Financial Economics*, *13*(2), 223–252.

Peress, J. (2014). The media and the diffusion of information in financial markets: Evidence from newspaper strikes. *The Journal of Finance*, *69*(5), 2007–2043.

Roll, R. (1988). R2. *The Journal of Finance*, *43*(3), 541–566.

Schwert, G. W. (1981). The adjustment of stock prices to information about inflation. *The Journal of Finance*, *36*(1), 15–29.

Shleifer, A., & Vishny, R. W. (1997). The limits of arbitrage. *The Journal of Finance*, *52*(1), 35–55.

Smales, L. A. (2014). Non-scheduled news arrival and high-frequency stock market dynamics: Evidence from the australian securities exchange. *Research in International Business and Finance*, *32*, 122–138.

Tetlock, P. C. (2007). Giving content to investor sentiment: The role of media in the stock market. *The Journal of Finance*, *62*(3), 1139–1168.

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

Tetlock, P. C. (2011). All the news thats fit to reprint: Do investors react to stale information? *Review of Financial Studies*, *24*(5), 1481–1512.

Tetlock, P. C., Saar-Tsechansky, M., & Macskassy, S. (2008). More than words: Quantifying language to measure firms fundamentals. *The Journal of Finance*, *63*(3), 1437–1467.

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, *5*(2), 207–232.

Veronesi, P. (1999). Stock market overreactions to bad news in good times: A rational expectations equilibrium model. *Review of Financial Studies*, *12*(5), 975–1007.

Wooldridge, J. M. (2013). *Introductory econometrics a modern approach*. South-Western, Cengage Learning.

A Appendix

A.1 Unconditional properties with group-specific means

As we see in Section 4.1, some companies have a strong weight in the average reaction to news articles. In fact, Equinor and Telenor accounts for 44% of the news articles. To underscore our results, we perform the same analysis with group-specific means, which give all companies similar weights, following Groß-Klußmann and Hautsch (2011).

Let n_s denote the number of news for stock s and let X_{sj} be the reaction in one of our variables of interest of stock s to news article j. To capture the average reaction of the n_n stocks with individual means, $\overline{X}_s = 1/N \sum_{j=1}^{n_s} X_{sj}$, we assume

$$\overline{X}_s = \mu + \varepsilon_s,\ \varepsilon_s \sim\ i.i.d.\ N(0, \sigma^2),\ s = 1, ..., n_n \tag{12}$$

Then, we draw inference based on the estimator for the mean, $\overline{\overline{X}} = 1/n_n \sum_{s=1}^{n_n}\ \overline{X}_s$ and the confidence intervals are given as $\overline{\overline{X}} \pm 2 * \frac{\widehat{\sigma}}{\sqrt{n_n}}$, where = $\widehat{\sigma}^2 = \boldsymbol{e}'\boldsymbol{e}/(n_n - 1)$.

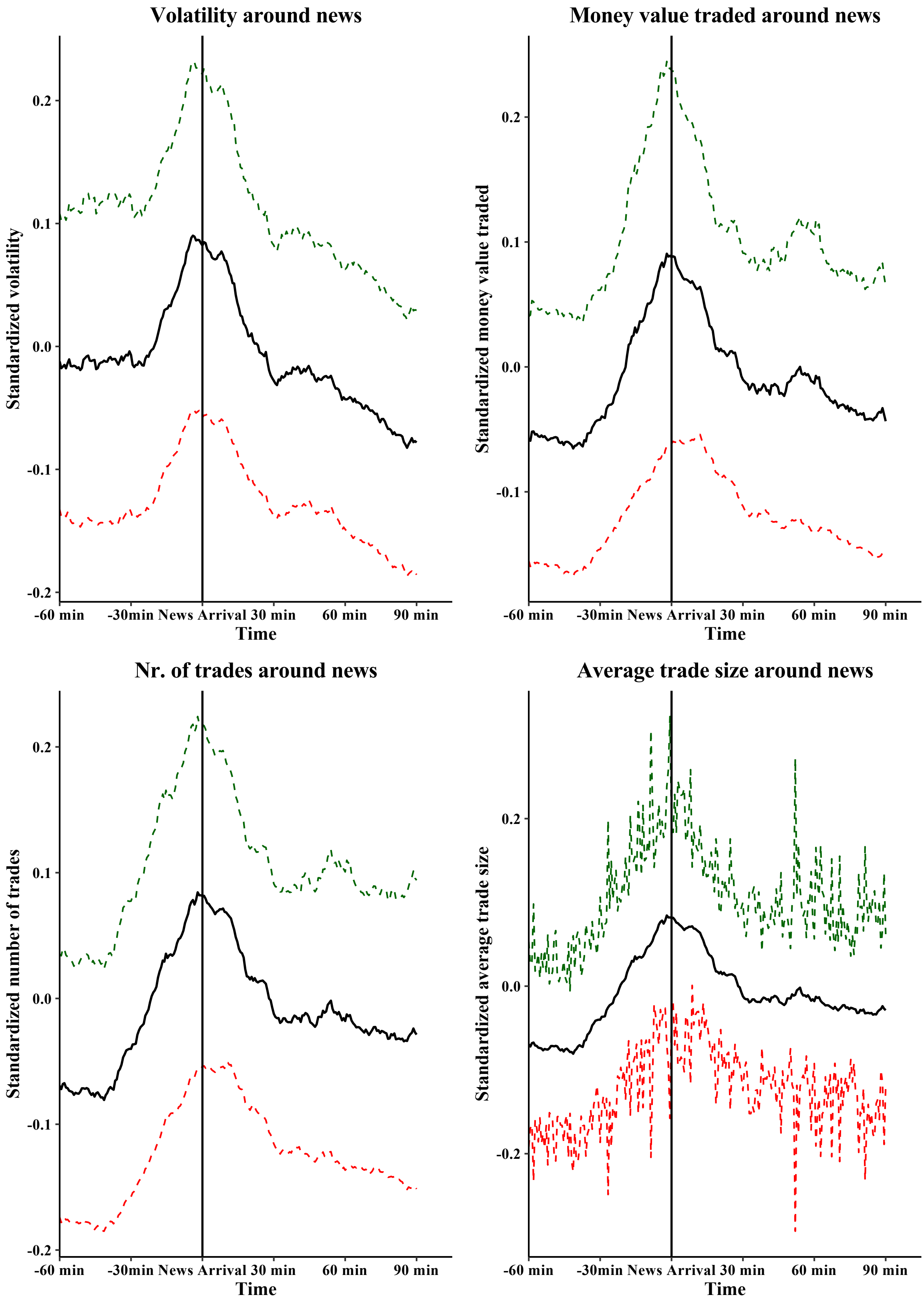

Figure 9: The average reaction in volatility, money value traded, average trade size and number of trades around news arrivals with group-specific means. The 95% confidence interval is illustrated with a green dotted line for the upper confidence bound and a red dotted line for the lower confidence bound.

A.2 VAR model with randomized news data

Table 8: Average VAR results with 5-minute intervals and randomized news data.

Dynamics		Money value	Volatility	Average trade	Nr. of trades
Constant	c	-0.006	-0.002	0.000	-0.005
		(0.012)	(0.014)	(0.014)	(0.012)
Money value	mv_{t-1}	0.246***	0.149***	0.120***	0.084**
		(0.039)	(0.043)	(0.044)	(0.038)
	mv_{t-2}	0.107***	-0.015	0.033	0.044
		(0.039)	(0.043)	(0.044)	(0.039)
Volatility	$vola_{t-1}$	0.016	0.097***	0.007	0.020
		(0.013)	(0.014)	(0.015)	(0.013)
	$vola_{t-2}$	-0.003	0.044***	-0.004	0.002
		(0.013)	(0.014)	(0.015)	(0.013)
Average Trade	ats_{t-1}	-0.077	-0.028	0.046**	-0.081
		(0.018)	(0.020)	(0.020)	(0.018)
	ats_{t-2}	-0.035	0.000	0.044**	-0.040
		(0.018)	(0.020)	(0.020)	(0.018)
Nr. of trades	n_{t-1}	0.008	0.049	-0.088	0.178***
		(0.036)	(0.039)	(0.040)	(0.035)
	n_{t-2}	0.019	0.017	-0.029	0.086**
		(0.036)	(0.040)	(0.041)	(0.036)
Dummy leads	d_{t+3}	0.008	0.064	-0.058	0.023
		(0.222)	(0.244)	(0.25)	(0.217)
	d_{t+2}	0.090	-0.145	0.075	0.027
		(0.222)	(0.244)	(0.25)	(0.217)
	d_{t+1}	0.191	0.105	-0.056	0.148
		(0.222)	(0.243)	(0.250)	(0.217)
News arrival	d_t	0.165	0.011	0.050	0.160
		(0.221)	(0.243)	(0.249)	(0.216)
Dummy lags	d_{t-1}	0.422	0.081	0.168	0.273
		(0.221)	(0.243)	(0.249)	(0.216)
	d_{t-2}	0.258	0.065	0.097	0.199
		(0.221)	(0.243)	(0.249)	(0.216)
	d_{t-3}	0.211	0.149	0.055	0.144
		(0.223)	(0.245)	(0.251)	(0.218)

Note: The table provides the average estimation results for the VAR models, as outlined in Equation 10. Estimates are provided for the dynamics of the endogenous variables, together with the exogenous news dummies. Reported coefficients are the average of the coefficients for each individual company with average standard errors given in the parentheses below. Significance is reported based on the average of the t-values. *** denotes the significance of the average coefficient estimates at the 1% level, ** at the 5% level, and * at the 10% level.

A.3 VAR model adjusted for weekday effects

Table 9: Average VAR results with 5-minute intervals adjusted for weekday effects.

Dynamics		Money value	Volatility	Average trade	Nr. of trades
Constant	c	-0.005	-0.010	0.006	-0.002
		(0.025)	(0.028)	(0.029)	(0.024)
Money value	mv_{t-1}	0.217***	0.131***	-0.057	-0.019
		(0.035)	(0.040)	(0.041)	(0.033)
	mv_{t-2}	0.126***	0.097**	0.027	0.081**
		(0.036)	(0.041)	(0.042)	(0.034)
Volatility	$vola_{t-1}$	0.085***	0.106***	0.025	0.084***
		(0.013)	(0.015)	(0.015)	(0.013)
	$vola_{t-2}$	0.002	0.031**	0.011	0.000
		(0.013)	(0.015)	(0.016)	(0.013)
Average Trade	ats_{t-1}	-0.018	-0.022	0.18***	0.010
		(0.015)	(0.017)	(0.017)	(0.014)
	ats_{t-2}	-0.041	-0.079	0.072***	-0.046
		(0.015)	(0.017)	(0.017)	(0.014)
Nr. of trades	n_{t-1}	0.111***	0.010	0.105**	0.372***
		(0.032)	(0.037)	(0.038)	(0.031)
	n_{t-2}	-0.032	-0.123	-0.063	0.024
		(0.034)	(0.039)	(0.040)	(0.032)
Dummy leads	d_{t+3}	-0.005	-0.005	-0.021	0.025
		(0.210)	(0.237)	(0.246)	(0.202)
	d_{t+2}	0.028	-0.159	0.096	-0.013
		(0.211)	(0.237)	(0.247)	(0.203)
	d_{t+1}	0.124	0.075	0.001	0.071
		(0.211)	(0.237)	(0.247)	(0.203)
News arrival	d_t	0.067	-0.003	0.073	0.079
		(0.211)	(0.237)	(0.247)	(0.203)
Dummy lags	d_{t-1}	0.487*	0.076	0.179	0.354
		(0.210)	(0.237)	(0.246)	(0.202)
	d_{t-2}	0.473*	0.483**	0.177	0.331
		(0.210)	(0.236)	(0.246)	(0.202)
	d_{t-3}	0.221	0.444	0.078	0.118
		(0.211)	(0.237)	(0.247)	(0.203)
Weekday		Yes	Yes	Yes	Yes

Note: The table provides the average estimation results for the VAR models, with dummy variables to control for weekday effects. Estimates are provided for the dynamics of the endogenous variables, together with the exogenous news dummies. Reported coefficients are the average of the coefficients for each individual company with average standard errors given in the parentheses below. Significance is reported based on the average of the t-values. *** denotes the significance of the average coefficient estimates at the 1% level, ** at the 5% level, and * at the 10% level.

A.4 Tables and figures

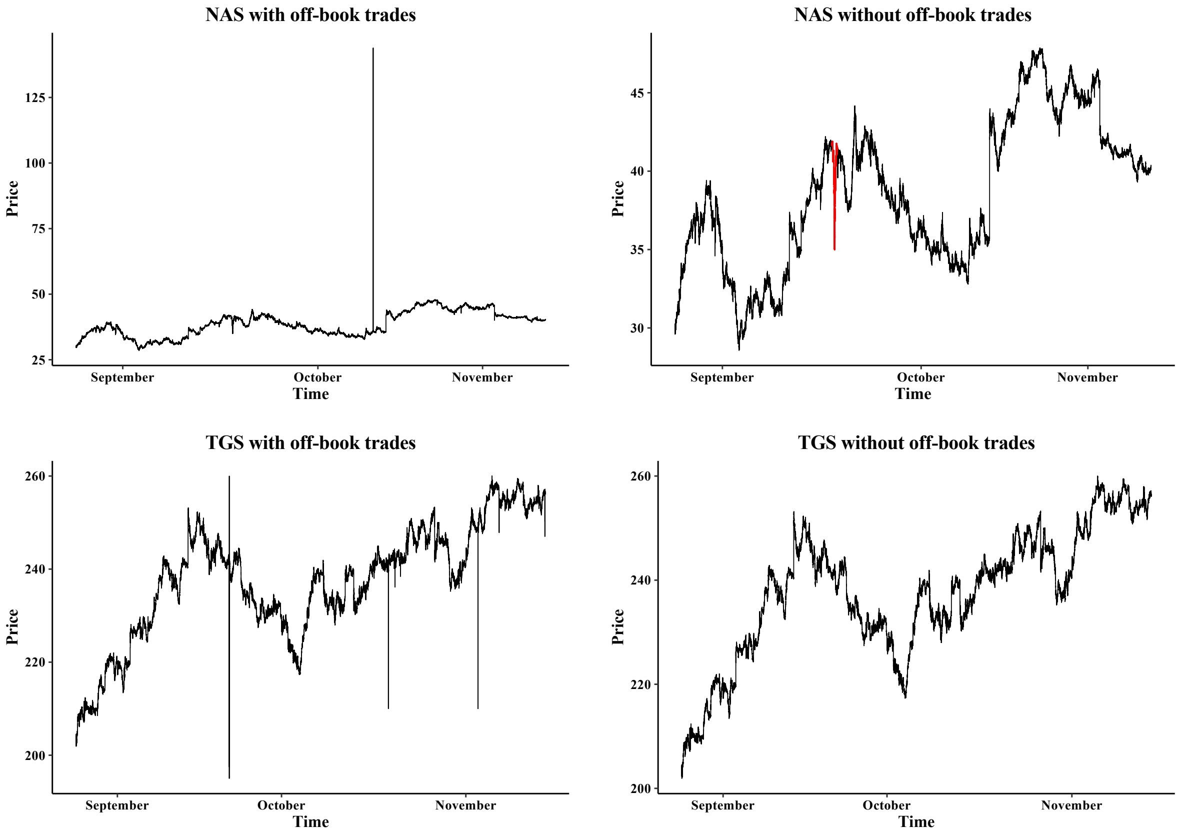

Figure 10: Time series of prices with and without off-book trades for Norwegian Air Shuttle and TGS-NOPEC Geophysical Company.

Publishers		
aksjelive.e24.no	benzinga.com	blogg.nordnet.no
bloomberg.com	bloombergquint.com	business.financialpost.com
businessinsider.com	businesstimes.com	businesswire.com
cerclefinance.com	cnbc.com	dn.no
e24.no	finance.yahoo.com	forbes.com
globalcapital.com	investing.com	londonstockexchange.com
marketwatch.com	newsweb.oslobors.no	reuters.com
seekingalpha.com	trader.di.se	

Table 10: Sample of every unique publisher.

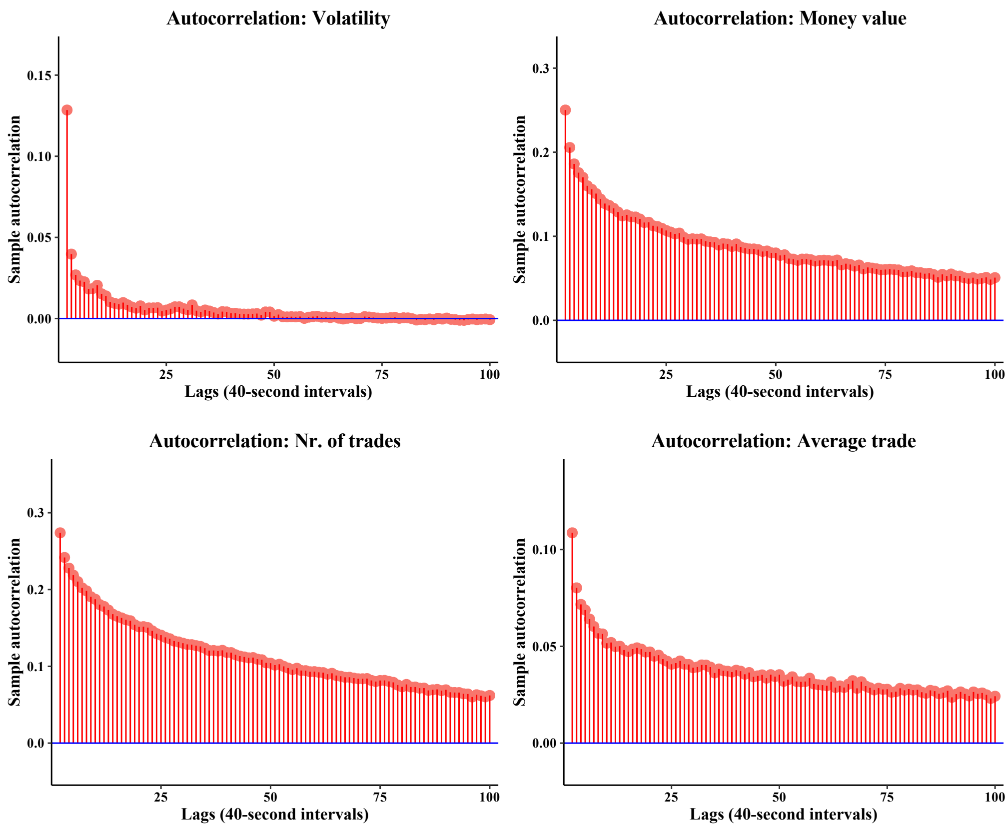

Figure 11: The average autocorrelation across every company.

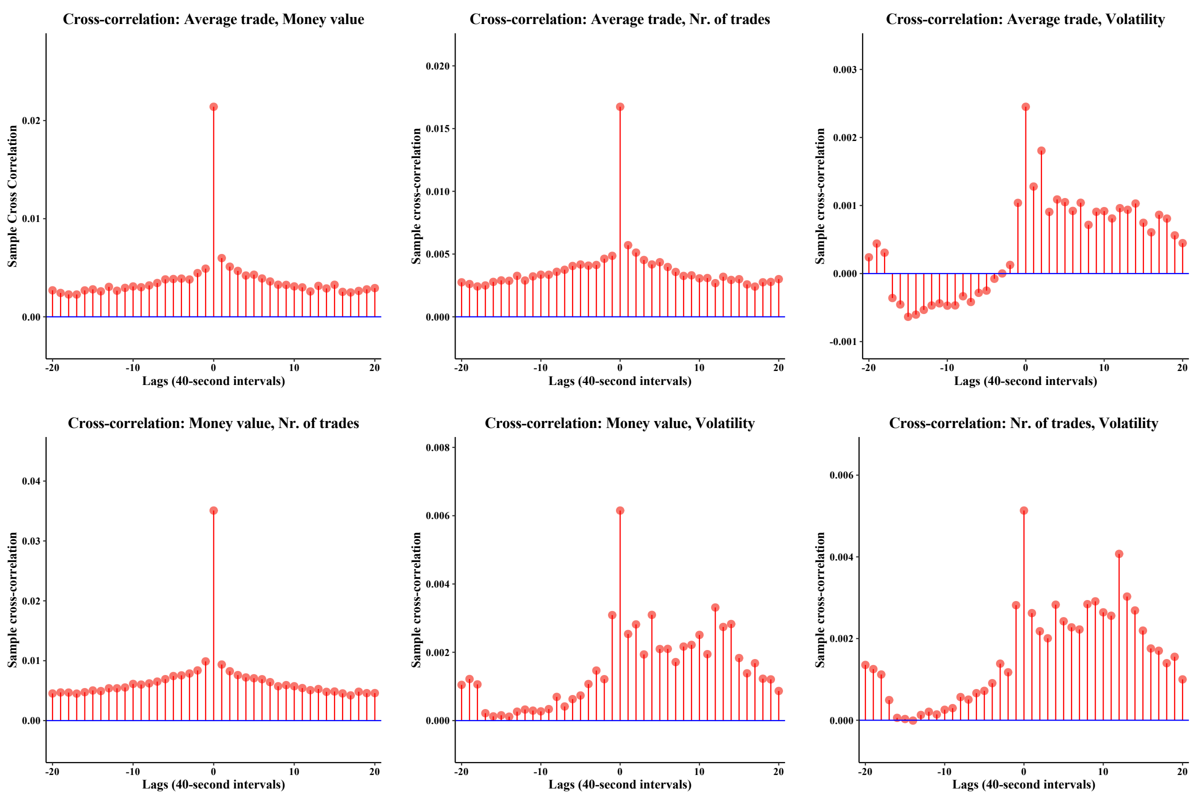

Figure 12: The average cross-correlation across every company.

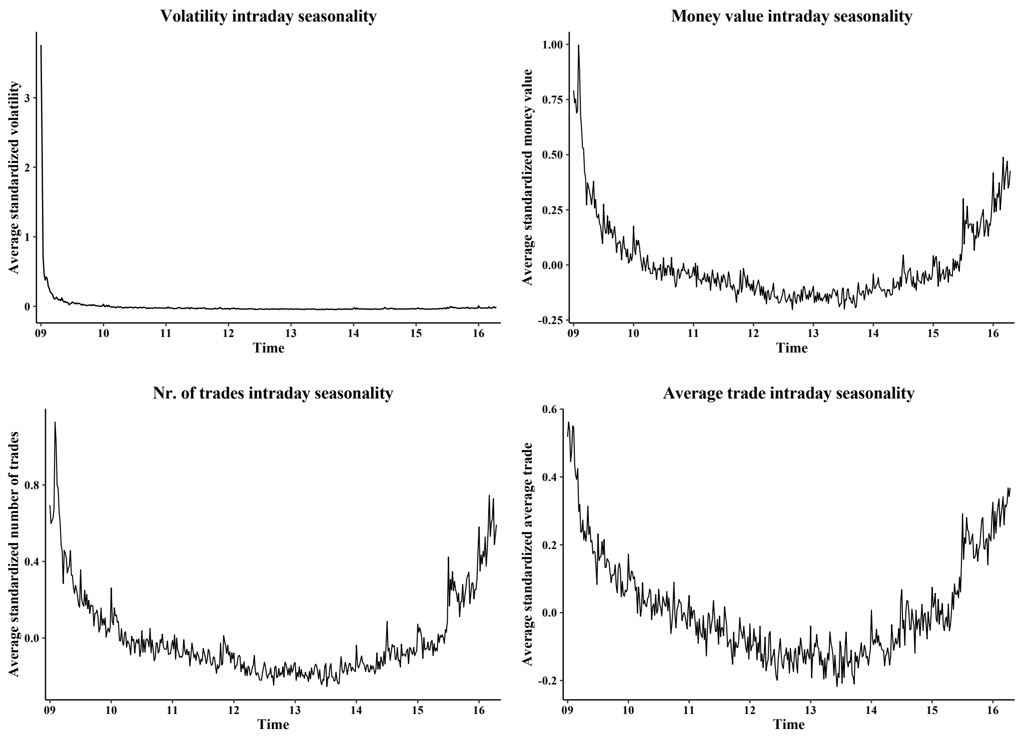

Figure 13: The average intraday seasonality patterns across every company.

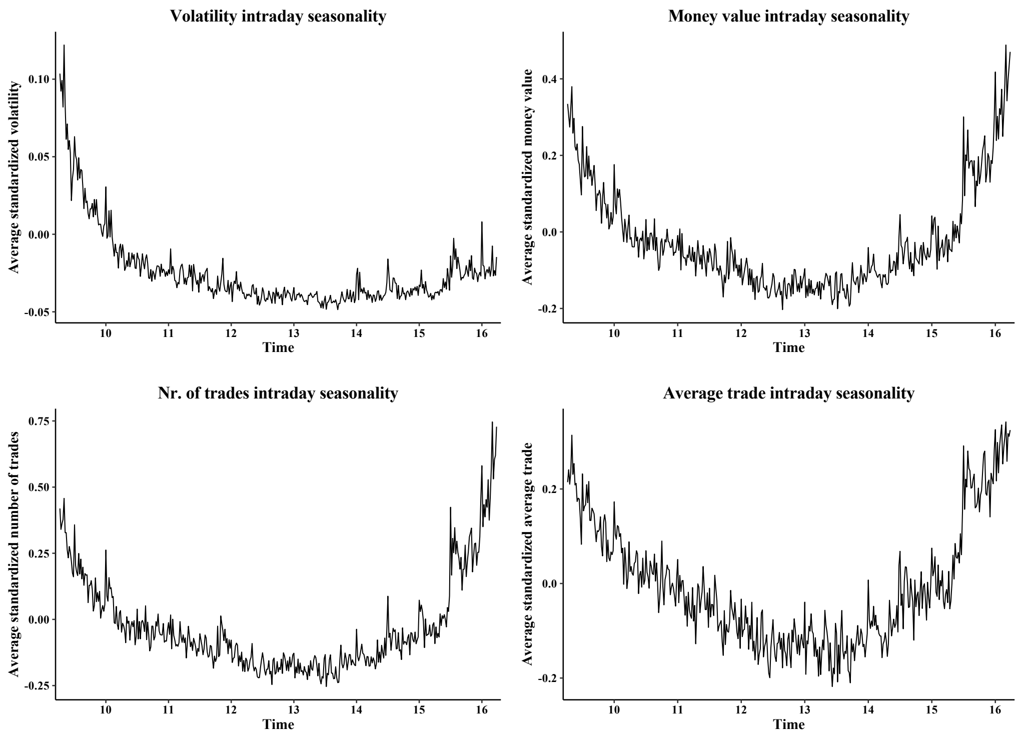

Figure 14: The average intraday seasonality patterns across every company without first and last 15 minutes of the trading day.