BI Norwegian Business School - campus Oslo

GRA 19502

Master Thesis

Component of continuous assessment: Thesis Master of Science
Final Master thesis – Counts 80% of total grade

Master Thesis

Navn: Henrik Woie, Sondre Lillebø Gundersen

Start: 02.03.2018 09.00

Finish: 03.09.2018 12.00

Master Thesis

News Impact by Announcement Frequency on the Oslo Stock Exchange

Authored by Sondre Lillebø Gundersen, and Henrik Olaus Woie

Supervisor:
Kjell Jørgensen

Hand-in-Date:
14. August 2018

Programs:
Master of Science in Finance, and
Master of Science in Business (Finance).

Abstract

This study looks at company-filing return variation with respect to firms' rate of market communication; the study spans all news-filings for every currently listed company on the Oslo Stock Exchange over the last twenty years. Specifically, it examines the variation of adjusted, abnormal returns by firms' twelve-month news announcement frequency. Analysis suggests the existence of an inversely related relationship between the frequency of announcements issued and news-event impact. The same result holds when controlling for choice of return adjustment model, measure of information flow, sample time-span, survivorship bias, and other biases. However, when analyzing a specific category of news – specifically, contract announcements – the relationship is conversely shown to be non-significant. We therefore conclude that a relationship does exist between announcement frequency and abnormal returns, but that the general relationship cannot be extrapolated to specific groupings of news within the aggregate sample.

Acknowledgements

We would like to thank our supervisor Kjell Jørgensen for his help and guidance throughout this process.

Table of Contents

Table of tables

Table of Figures

1. Introduction

Most listed firms on the Oslo stock exchange (OSE) issue announcements several times per year, sometimes per week. The nature of the issuances ranges from wildly positive, as in the case of winning lucrative contracts, to wildly negative, as in the case of initiating insolvency procedures; and investors respond accordingly. This paper aims to establish whether there exists a relationship between the rate of firms' news announcement and returns. The study spans all currently listed companies on the OSE, and all news events issued in the last twenty years. Specifically, we hypothesize that return impact ought to be inversely related to the announcement frequency of firms, i.e., that firms with less frequent market communication leading up to an event might experience stronger investor reactions relative to firms that communicate more frequently. Investor reactions in this setting refers to the effect on stock returns measured in abnormal returns, and the frequency of announcements is defined as the number of announcements leading up to an event, over a twelve-month period. The underlying intuition and motivation of the research question is in part based on the assumption that companies which inform investors more frequently will be priced closer to their true intrinsic value at any given time, which over time should lead to smaller price change per announcement, relative to infrequently communicating firms, given they communicate the same aggregate information over a fixed time horizon. If this is the case, we expect to see it manifested in an inverse relationship between announcement rate and impact as measured in abnormal returns. The paper draws heavily on academic literature with a primary focus on classical event study literature and specifically event studies focusing on the impact of news on stock returns. The exact application of the study however, is to our knowledge not extensively covered elsewhere. To test our hypothesis, we apply classical event study methodology and conduct event studies for each individual news issue in our sample. Abnormal returns and other event-specific characteristics are computed for each individual event before all observations are aggregated across companies to form our sample. We use linear regression to estimate the impact of announcement frequency on abnormal returns. In our analysis we find that a general relationship does seem to exist for the mass of news issued by currently listed companies on the OSE over the last twenty years. This result holds across methodological variations of the measure for information flow; holds in the first half and second half of the sample period independently; it holds for three separate return-adjustment models; and they hold when controlling for a frequency bias shown to exist in the sample. Conversely, results indicate that the general relationship found does not hold for isolated subcategories of news, like the category of contract announcements. This result seems to indicate that the general relationship found in the main analysis cannot be extrapolated onto all subcategories of news. The remainder of the paper is organized in nine main sections, starting with the introduction; a review of the theory and literature; the research methodology; a look at the data; the main analysis; a section dedicated to robustness checks; a secondary sub-analysis looking specifically at the category of contract announcements; the conclusion; and a short section dedicated to discussing study shortcomings and general retrospection.

1.1 Research Question

The economic intuition for our hypothesis is - in part - laid out in Fama's *Efficient Capital Markets* (1970), where Fama posited that the expected price of a security should equal the current price factored by the expected one-period percentage return conditional on the information set, Φ_t which should be fully reflected in the stock price. It follows that the return, $x_{j,t}$,equals the realized price $p_{j,t+1}$ minus the expected price as shown below.

$$x_{j,t} = p_{j,t+1} - E(p_{j,t+1}|\Phi_t)$$

Where the expected return $\mathrm{E}(\tilde{x}_{j,t+1}|\Phi_t) = 0$, as the information set Φ_t ought to be fully reflected in the price, meaning trading on the information should not yield abnormal returns for investors. This largely sets the basis for the efficient-market hypothesis and a large literature has been dedicated to testing its validity. This paper - *rather* than testing market efficiency by looking at whether abnormal returns can be made in equilibrium – will analyze whether certain characteristics of Φ_t, has a measurable impact on the returns of news-filings on the OSE. For practical purposes we restrict the scope of the paper to news-filings issued through the exchange only, and therefore define our information set φ_t as the sum of prior news-filings in a given security:

$$\varphi_{i,t} = \sum_{n=1}^{N} \kappa_{n,i}$$

Where κ_i represents news-filing *n* for company *i*. Simply put, this paper looks at whether observable differences in abnormal-returns exists for new-information-issues, news-events, with respect to previous information-flow.

2. Theory and Literature Review

The subset of literature most important to this paper falls under the domain of event study literature. In particular, the literature on reactions to news events, and reactions to earnings announcements is useful. The literature on event studies was largely popularized by Fama, Fisher, Jensen, and Roll (1969) and has since grown to become a common method for measuring the impact of an event. Event studies are designed to measure the impact of an event by measuring the generated abnormal returns. At the same time - as highlighted by MacKinlay (1997) and Brooks (2014) among others - event studies are often considered to be tests of market efficiency: given rationality in the marketplace, there should be an immediate reaction to the event on the announcement date and no further reaction on subsequent trading days, given no new significant information. Though much of the literature early on looked at events such as dividend initiation, stock splits, acquisitions, or security offerings (Barber & Lyon 1997), event studies can be applied in most cases where there is an easily definable event; in our case, any filing from the OSE will be classified as an event. Traditional event study methodology will be covered more extensively in the methodology section, but to a large extent the paper stays true to the methodology laid out by MacKinlay (1997) and Brooks (2014). The primary differences in the methodology of this paper is that we suspect markets require days rather than weeks to absorb the impact of the average news event and we therefore establish a shorter than usual event-window relative to the fundamental literature. For the literature on reactions to news, there is to our knowledge no available literature on the effect of news with respect to time and frequency, and we instead draw inspiration from the large literature on the post-earnings announcement drift. Two articles of particular interest to us was Zhang (2006) and Hirshleifer et. al (2009). Zhang refines the concept of post-earnings announcement drift by focusing on information uncertainty and the ambiguity of new information that investors face. His evidence supports the phenomenon of post-earnings announcement drift and finds that firms with higher information uncertainty has greater drift. He points to two potential sources of information uncertainty, the first being volatility in firms' fundamentals, and the second being poor information. Hirshleifer et. al proposes the investor distraction hypothesis, suggesting that limited investor attention may cause market under-reactions. They test this by looking at the level of overreaction and the intensity of news flow measured by the daily number of announcements. Results showed that not only are there definite signs of drift in their study, but that extraneous news seems to amplify the effect. Specifically, they find that the inter-decile spread of announcement-period abnormal returns between firms with high and low earnings surprises being 7.02% for low-news days and 5.81% for high-news days. The results are highly relevant to our study since we hypothesize that there is an inverse relationship between the rate of announcements and absolute levels of impact. The main difference of course is that Hirshleifer et. al (2009) looks at the effect of news as a function of the daily frequency of total news, while we examine the frequency of firm-specific news over a rolling window of time.

3. Methodology

This section will be dedicated to presenting the methods and tests we utilize and our underlying motivation for doing so. The section will be split into two subsections: event study methodology and regression-specific methodology & design.

Event Study Methodology

To be able to test the hypothesis we have presented, we need a method for measuring event impact. For this purpose, we implement classical event study literature as presented by Fama (1970; 1991), MacKinlay (1997), Armitage (1995), and Kothari and Warner (2004) among others. The method consists of defining an event window, computing a measure of expected return, and computing adjusted returns for the event window as a measure for impact.

3.1 Defining the Event

In the classical event study literature, an event is typically related to news released by the financial press or news released by companies (Peterson 1989). We focus on the latter, and consider all news released via *Newsweb* - the OSE website for firm specific news – to be relevant; the database contains data in the interval 1998 – 2018 for all currently listed companies. The day of issue serves as the event day and returns from that day is attributed to the computation of event returns with one notable exception: if announcements are made by a company after trading hours have concluded, we treat those events as if they had been announced prior to trading commencement the following day. No data is otherwise changed, and the difference in treatment for post-market issuances refers to the attribution of returns only.

3.2 The Event Window

Our event window spans two full days of trading. It is typical to include additional time after the event to allow time for the full effect of the event to be absorbed into the stock price (MacKinlay 1997). This could mean including extra minutes, or extra months, depending on what is being studied. For this study, what is optimal likely depends on the event, as the time required by investors to assimilate new information will depend on the ambiguity and nature of the filing. For instance, one would imagine that the market could price a $10 million grant receival more quickly than news of a company entering a new market, because the former has fewer parameters. While authors like McWilliams and Siegel (1997) advocate for shorter event-windows to mitigate the risk of including confounding events, estimating an appropriate event-window for the study is challenging, as we expect the lower-bound of information processing to have changed during the twenty-year sample. Empirical evidence however, seems to suggest the notion that observations going back to the start of the sample ought to be priced in minutes to hours rather than days.[1] We therefore consider a two-day post-event window to be appropriate and conservative.

3.3 The Estimation Window

Our estimation window is set at 126 days, or approximately six months' worth of trading days. Brooks (2014) highlighting the trade-off between increasing the precision of parameter estimation and raising the likelihood of structural breaks being included in the window, while Armitage (1995) suggests using 100-300 days; in aggregate we find 126 days to be reasonable.

3.4 Abnormal Returns & Adjustment Model Selection

To estimate the impact of events we compute abnormal returns using a *market model*. We do this by adjusting individual firms' returns, $R_{i\tau}$, in the event window by the expected return for the individual security, i.e., we calculate the returns above a measure of the expected return, estimated for the individual firm *i*, for each day τ, in the event window:

$$AR_{i\tau} = R_{i\tau} - E(R_{i\tau}) \tag{1}$$

where $AR_{i\tau}$, $R_{i\tau}$, and $E(R_{i\tau})$ are the abnormal returns, actual returns and normal returns respectively. The expected returns can be calculated using different models, and for this paper we utilize three: The *Market Model*, the *Constant Mean*

[1] Ryngeart and Netter (1990); Dann, Mayers, and Raab (1977); Mitchell and Netter (1989); and, Ederington and Lee (1993).

Model, and the *Fama-French 3-Factor Model*. While Mackinlay (1997) highlights the benefit of using factor models over a constant mean model as the variation in the abnormal returns due to variation of market factors can be explained away, he also points out that the benefit from a two-day study like ours is less clear. We do not expect the factor models to perform much better than the simple constant mean model on such a short horizon, as supported by Marshall et. al. (2017), and therefore do not have strong preferences for choice of model. As a solution we employ all three - the market model for our main analysis and the constant mean- and factor model for robustness checks in the robustness analysis section. For firm *i* in the sample, the models are described:

$$Constant\ Mean: \quad R_{i\tau} = \mu_i + \varepsilon_{i\tau} \tag{2}$$

$$Market: \quad R_{i\tau} = \alpha_i + \beta_i R_{m\tau} + \varepsilon_{i\tau} \tag{3}$$

$$Fama-French: \quad R_{i\tau} = a_i + b_i R_{m\tau} + s_i SMB_\tau + h_i HML_\tau + \varepsilon_{i\tau} \tag{4}$$

Where for the market model $R_{i\tau}$ and $R_{m\tau}$ are returns on firm *i* and the market portfolio *m* respectively, for day τ, and $\varepsilon_{i\tau}$ is the error term. The market portfolio is the return of the market, represented by the *Oslo Stock Exchange All-Share Index* (OSEAX) adjusted by the risk-free rate. The market model and the Fama French-model are estimated by ordinary least squares (OLS) regression.

3.5 Cumulative Abnormal Returns

The cumulative abnormal return of firm *i* for each event is the sum of abnormal returns in the event window, and is defined as:

$$CAR_i[\tau_0, \tau_1] = \sum_{\tau_0}^{\tau_1} \mathrm{AR}_{i,\tau} \tag{5}$$

Where τ_0 is the time of the event, and τ_1 signifies the end of the two-day post-event window. To account for the compounding effect investors experience, we also compute buy-and-hold returns, *BHARs*, for the event window the following way:

$$BHAR_i[\tau_0, \tau_1] = \prod_{\tau_0}^{\tau_1} (1 + R_{i\tau}) - \prod_{\tau_0}^{\tau_1} \left(1 + R_{m,\tau}\right) \tag{6}$$

While *BHAR* seems most appropriate for longer studies, we will use *CARs* as our primary measure of impact. For data descriptive purposes, and purpose of comparison, we will use one additional measure, namely aggregated CARs, *CAAR*s. CAARs are defined:

$$CAAR(\tau_0, \tau_1) = \frac{1}{N}\sum_{i=1}^{N} CAR_i(\tau_0, \tau_1) \tag{7}$$

Regression Analysis

Following the implementation of event study procedures and sample generation, regression analysis enables us to examine whether relationships between a given variable and one or more other variables exist (Brooks, 2014). This of course cuts to the core of our research question which is whether a relationship exists between *Frequency* and abnormal returns.

3.6 The Regression Model

Following the computation and aggregation of abnormal returns from the event study part of the paper, we want to analyze whether the frequency of announcements preceding an event can explain the variation in announcement returns, *CAR*[0,1]. We use the following regression as our baseline model for testing:

$$|CAR_{[0,1]}| = \alpha + \beta_1 Frequency + \beta_2 Positive + \beta_3 Frequency\ x\ Positive + \varepsilon \tag{8}$$

where α is the intercept, |C*AR*[0,1]|is the 2-day absolute cumulative abnormal return for every event in the sample, *Frequency* is the 12-month event-specific frequency measure, *Positive* is an indicator variable that is equal to one when CAR[0,1] > 0 and zero otherwise, *Frequency x Positive* is *Positive* multiplied by *Frequency*, and ε is a random error term. The alternative hypothesis to be tested is whether frequency has an impact on abnormal returns, i.e., $\beta_1 \neq 0$. Seeing that the effect of announcement frequency, captured through β_1, could potentially be asymmetrical with respect to positive and negative announcements, we use the coefficient β_2 to capture the differences in positive and negative intercepts, and the coefficient β_3 to capture slope asymmetries. All regressions model abnormal returns as a dependent variable in absolute terms. The transformation of data to absolute terms was done to enable more meaningful model interpretations. As the original data is two-sided, we expect positive and negative events to largely cancel out, rendering the model meaningless for purposes of proving a general relationship between impact and frequency. By regressing in absolute terms, we effectively split the model in two and are able to discern the general relationship of frequency and impact as well as the differences between the halves. In short, the change leaves us with a model measuring impact rather than the aggregate differences in positive and negative news composition.

Once the analysis of the baseline model is concluded we introduce a range of appropriate control variables covered in section 4.2 to address the issue of omitted variable bias which could explain some of the variation in announcement returns. For the expanded regression including control variables we use the following regression specification:

$$|CAR|_\tau = \alpha + \beta_1 Frequency_\tau + \beta_2 Positive_\tau + \beta_3 Frequency\ x\ Positive_\tau + \sum_{i=1}^{n} \varphi_i X_{i\tau} + \varepsilon_\tau \quad (9)$$

where the major change relative to the baseline model is the inclusion of X, an Nx1 vector of control variables and φ which is an 1xN vector of coefficients. The alternative hypothesis to be tested remains whether frequency has an impact on announcement returns, |*CAR*[0,1]|, i.e., $\beta 1 \neq 0$. For notational ease, we refer to |*CAR*[0,1]| simply as *CAR*[0,1] after introducing the baseline regression model in section five.

4. Data

This section is dedicated to explaining how we collect data, what data is included, and how it is adapted. The section is organized in three parts: the first part explains the data collection procedure, the second part defines relevant variables, and the third discusses data-specific methodological considerations.

4.1 Data Collection

The study makes use of data from several sources, but we largely gather and generate data from three main ones: firm-specific data, factor data, and news data. First, the category of firm-specific data refers to data collected using *Datastream,* the financial and economic research data database from Thomson Reuters. It includes variables such as closing price; opening price; intraday-high-, and intraday-low price; market capitalization; shares outstanding; turnover; trading volume; listing dates; and industry affiliation. All data is collected as daily data. Second, as recommended by our thesis advisor we collected factor data required for the market model and factor model abnormal return calculations as supplied by Ødegaard (2018). We were happy to find that complete Fama French 3-factor portfolios calculated on Norwegian data, as well as index data and data for the Norwegian risk-free rate was maintained and made available. As mentioned, we use the OSEAX as our market index for return calculations. Third and last, we transcribed news data from *Newsweb*, the official outlet of the OSE using programming logic. The sample spans all events of all currently listed companies, as made available on the website.

4.2 Variables

The sample of data we use for regression analysis contains roughly fifty distinct variable-columns spanning almost thirty thousand rows of data. Variables can largely be separated into the following categories: firm characteristics, news specific variables, return data and abnormal return data. Further variables such as price data are left out after the event studies and sample aggregation is finished. This subsection looks to define each relevant variable and its functions in the coming analysis.

Our primary exogenous variable is our measure for information flow, *Frequency*. *Frequency* is the number for the rate of announcements within a company - in the last twelve months - at the time of a news event. It is news-specific and calculated using a rolling estimation-window. In addition to the twelve-month estimation-window *Frequency* we compute *Frequency-3m* which uses a three-month window, *Frequency-6m* which uses a six-month window, and *Interval* which we later use to test the robustness of *Frequency* as our measure of information flow in the main analysis. *Interval* is defined as the number of days between current- and closest preceding event and serves as an alternative measure of information flow. From the event study procedures, we retain mostly abnormal return data to serve as our endogenous variable throughout analysis, and for robustness checks in section 6. The main variables include AR_0, AR_1, $CAR_{[0,1]}$ calculated using three separate models – the mean-, market-, and factor model as covered in section 3.4. For news specific data we record the time of the event, the headline, the ticker of the company, and the date of issuance. News examples are illustrated below.

Index	Time	Ticker	News
11.11.1998	15:59:03	AFG	MELDEPLIKTIG HANDEL
14.01.1999	12:17:36	AFG	FLAGGING/MELDEPL. HANDEL
…	…	…	…
14.12.2012	08:46:16	MHG	Share purchase programme for employees in Marine Harvest/Purchase of own shares
17.12.2012	09:01:05	MHG	Mandatory notification of trade
…	…	…	…
25.09.2017	15:20:57	ZAL	Acquisition of ROC Global Solution Consulting Ltd.
10.10.2017	08:35:22	ZAL	Zalaris ASA (ZAL): Acquisition of the remaining 2.68% of the shares in Sumarum AG

Figure 1: News Data

Collecting the data ourselves saved us time, but in turn meant settling for poorer, less detailed data than could have been purchased directly from the exchange. In other words, while we are able to obtain information on the time of the event; the weekday of the event; the source of the filing; and the headline of the event, we find that the lack of further information such as categorization of events and filing contents puts limitations of what level of analysis is possible. Below follows an overview of the distribution of news over time.

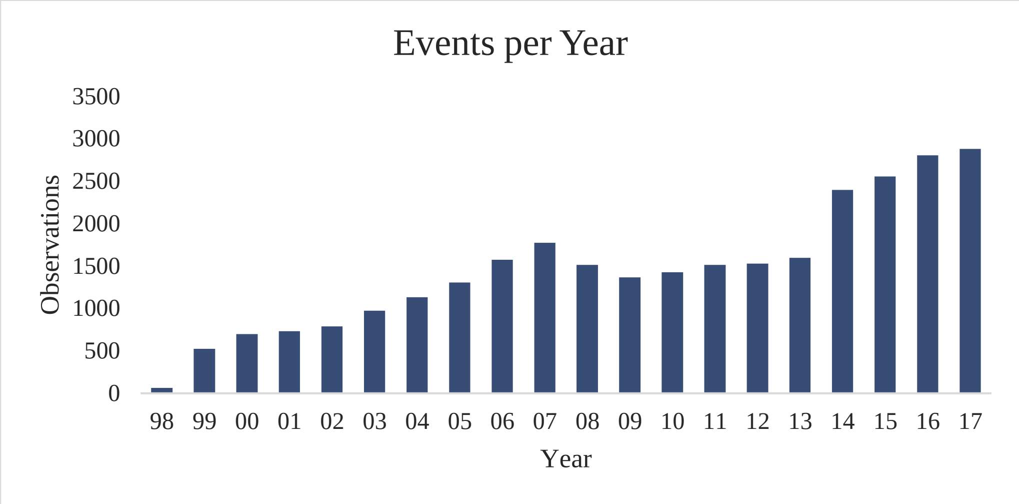


Figure 2: News Events Per Year

In addition to the raw event and return data, we also collect a large volume of data to serve as control variables. The reason for this is simply that, if a relationship is shown to exists between announcement frequency and abnormal returns in our baseline regression, we want to try and explain it using control variables in an extended regression. As pointed out by Hirshleifer et. al. (2009), several papers have found relationships between impact and proxies for investor inattention which may help us select reasonable control variables for our purposes. For instance, issuing news during non-trading hours is found to decrease impact[2] – in our case we include this as a factor by using a set of four indicator variables: *Pre-Market*, *First-Half*, *Second-Half*, and *Post-Market*. Pre-market is defined to be prior to trading commencement at 09:00 Norwegian time; the first-half is defined as between trading commencement and 12:15; the second-half is defined as between 12:15 to 16:25; and post-market is defined as 16:25 to midnight. See descriptive statistics for *Time of Day* variables below.

Table I
Descriptive Statistics - Time of Day Variables

	Distribution	Observations
Pre Market	41.9 %	12 217
Post Market	8.5 %	2 487
First Half	24.5 %	7 156
Second Half	25.0 %	7 290

[2] Francis, Pagach, and Spehan (1992); Bagnoli, Clement, and Watts (2005)

As shown in the table, the majority of events are issued before the market opens and a strong minority after closing. Another documented effect includes what is known as the Friday effect.[3] Rather than adding a single indicator variable for Fridays, we include a set of indicator variables representing each day of the week as shown below.

Table II
Descriptive Statistics - Weekday Variables

	Distribution	Observations
Monday	17.4 %	5 084
Tuesday	19.2 %	5 608
Wednesday	20.9 %	6 097
Thursday	21.7 %	6 332
Friday	20.7 %	6 029

For weekdays we observe a slight majority of events issued on Thursdays, and a lower number of announcements made on Mondays. A third relationship is that of low trading volume decreasing impact.[4] For this relationship we include a measure for turnover over value that we term *Relative Turnover.* Next, as a proxy for firm size, we use *Value* – the market capitalization of the firm at the time of the event - as we suspect the announcement frequency of companies will generally increase as firms increase in size. We also include a proxy for company maturity, as measured in total years listed at the time of the event, *Years Listed*. Summary statistics of these three variables are shown below; note that *Relative Turnover* is displayed in percentage points and *Value* is displayed in billions.

Table III
Descriptive Statistics - *Years Listed*, *Relative Turnover*, and *Value*

	Observations	Mean	Std	Min	25%	50%	75%	Max
Years Listed	29150	8.10	5.16	0.67	3.71	7.25	11.66	19.96
Relative Turnover	29150	10.00	115.09	0.00	0.03	0.15	0.51	4639.53
Value	29150	15.30	52.26	0.00	0.49	2.05	7.57	628.80

[3] DellaVigna and Pollet (2009)
[4] Hou, Peng and Xiong (2008)

From the table we can infer that the median company on the exchange was listed in 2010, has a relative turnover of 0.15% on the day of the event, and is worth roughly two billion Norwegian kroners. Next, we include indicator variables for industry affiliation, to address the endogeneity of firm's announcement policy (Kothari 2004), i.e., the problem of variations in reporting standards and the ability of companies in certain industries to determine when and if an announcement will be made. In sum, the industries vector includes 11 distinct industry indicator variables. Descriptive statistics of industries are shown below.

Table IV
Descriptive Statistics - Industry Variables

Variable	Distribution	Observations
I01 - OSE50 Telecommunication Services	1.5 %	442
I02 - OSE15 Materials	5.0 %	1451
I03 - OSE35 Health Care	4.2 %	1230
I04 - OSE55 Utilities	1.1 %	312
I05 - OSE60 Real Estate	2.9 %	855
I06 - OSE10 Energy	31.5 %	9187
I07 - OSE30 Consumer Staples	5.3 %	1547
I08 - OSE40 Financials	8.4 %	2450
I09 - OSE20 Industrials	22.3 %	6495
I10 - OSE45 Information Technology	14.4 %	4210
I11 - OSE25 Consumer Discretionary	3.3 %	971

The majority of announcements belong to the Energy sector and Industrials, as one might expect, with the two sectors accounting for more observations than all other categories combined, with 53.8% of the sample.

4.3 Data Considerations

In addition to the general introduction of variables, we feel there is a need to include a short section on specific considerations and methodological choices we have made because of the data and its structure. This section therefore, addresses data-specific issues.

Overlapping Events & Noise

As explained in section 3.2, we implement a two-day post-event window for our abnormal return calculations. What this section does not make clear is the issue of managing overlapping events within a single security. We are lucky to have a high number of events to analyze, but also have a large number of overlapping events in the sample that – if left as they are – will create a sample bias. Simply explained, if we have three events in the same day, and analyze them individually, we effectively triple-weigh these observations. To control for this, we implement a two-stage solution: First, we eliminate events that we consider to be unimportant, i.e., plain noise that we do not expect should earn any abnormal returns because they do not contain new information. These events primarily include routine announcements relating to quarterly presentations, invitations to annual general meetings, and filings regarding the ordinary financial calendar (for a full overview of the list of terms excluded, see appendix A1). Thereafter we eliminate all remaining overlapping events by excluding them from the sample altogether. In sum, this reduces the aggregate sample from roughly 123,000 events to 29,000, where roughly 20,000 are removed because of overlaps, and the remaining 74,000 are noise. Of course, by removing the overweighting bias, we introduce another bias in the sense that the removal of overlapping events likely affects more high-frequency issuers. This problem is analyzed and controlled for in section 6.5.

Listing Effects

A well-documented market phenomenon is listing effects, as documented by Kadlec & McConnell (1994) among others. Said simply, newly listed companies tend to experience strong positive returns during the initial period after listing. To eliminate any listing effects from our estimation windows, we eliminate data from the first three months of trading for newly listed securities.

5. Analysis

In the coming section we will cover a concise retelling of our hypothesis in addition to the analysis and relevant result of our work. The hypothesis to be tested is whether news events experience different reactions given differences in information flow leading up to events, where information flow is limited to past news events issued by the exchange. Specifically, we have hypothesized that lower information flow ought to lead to increased impact relative to larger information flow; impact meaning greater *absolute* abnormal returns. Simply put, we wish to test whether less news equates to stronger investor reactions and vice versa. For the remainder of the paper we will refer to firms with lower information flow as firms with lower *frequency* (referring to the lower frequency of announcements over a 12-month rolling window at the time of the event), and will refer to events with positive and negative CARs as simply positive and negative events. The analysis will be divided into two main sections: first we introduce the data, visually investigate differences in impact across an aggregate dataset of positive and negative events, and use regression analysis to model the effects of frequency on announcement returns. In the regression we use absolute returns and account for differences in positive and negative news by implementing a set of dummy variables. Second, we address the potential problem of omitted variables, and look to explain away any effects of frequency on returns by including control variables associated with company size; industry effects; time of day effects; weekday effects; and maturity effects.

5.1 The Baseline Model

While the assumption of increased impact given less previous information flow, or market communication by a company seems sensible to us, this may not be the case. To assert whether a relationship seems to be present in the first place, we start the section with a general look at the data. For purposes of comparison and testing we split the sample into quartiles based on *Frequency*. Data spans all news events on the OSE, in the last twenty years - starting at the inception of *Newsweb* in 1998. Descriptive statistics of the data are displayed below.

Table V
Descriptive Statistics – Frequency Quartiles

Panel A: Distribution of Number of Announcements								
	Observations	Mean	Std	Min	25%	Median	75%	Max
CAR[0,1]	29150	0.0043	0.0698	-0.7962	-0.0177	0.0009	0.022	2.9929
Frequency	29150	25.31	14.23	0	16	23	32	151

Panel B: Quartile CAR [0,1] Sample Characteristics								
	Observations	Mean	Std	Min	25%	Median	75%	Max
Q1	8 230	0.0060	0.0859	-0.7452	-0.0187	0.0009	0.0239	2.9929
Q2	6 713	0.0036	0.0693	-0.7141	-0.0190	0.0010	0.0233	1.3475
Q3	7 165	0.0045	0.0603	-0.7962	-0.0169	0.0008	0.0221	1.0130
Q4	7 042	0.0027	0.0573	-0.7607	-0.0163	0.0008	0.0197	0.8801

In aggregate, the dataset contains almost thirty thousand observations, with an overall mean CAR above zero at 0.43%; conversely, the median is much closer to zero at 0.09%, indicating a large positive variability and the presence of outliers.[5] There is a difference in standard deviation of 2.86 percentage points between the first and last quartile, and though trends in the remaining data is not evident, there does seem to be a clear anticorrelation of both returns-, and variability of returns to *frequency* - as hypothesized. As touched upon in the introduction, the differences in standard deviation of quartiles are particularly interesting to note, as we are currently working with a two-sided dataset, meaning the set contains both negative and positive events at all levels of frequency. Given a uniform increase in impact on either side of the distribution at lower levels of frequency relative to higher ones - under the assumption of rough symmetry of positive and negative events - we expect volatility to increase with the numbering of quartiles. In other words, we may here use the difference in volatility as an indication of the relationship we are looking for. To validate that the perceived difference has statistical validity, we run a right-tailed F-test on differences in quartile variances and find that the test indicates σ^2_{Q1} to be significantly different and higher relative to all other quartiles above it.[6] In summary, we find that - at first glance - the data seems to indicate that a relationship exists between *frequency* and abnormal

[5] Mean significantly different from zero at the 1% level. See Appendix A2 for results.
[6] The test of Q1 to Q4 returns an F-statistic of 2.25 over a critical value of 1.04. See Appendix A3 for complete test statistics.

returns in the direction hypothesized. Of course, this is far from conclusive, and we have not yet started to account for a large range of factors that might impact the strict comparability of baskets.

A fundamental issue of our regression design highlighted in Table V is that positive and negative events largely cancel each other out, and we get means close to zero. The implications of this is that a regression model on raw return data will likely not model the impact of news, but rather the asymmetry of positive and negative events; coefficient significance will not be interpreted meaningfully, because whether frequency truly explains changes in impact or not, the slope of frequency as an exogenous variable in a regression analysis will not tell us much if increased positive impact primarily results in increased variability, and not in meaningful changes of aggregate returns. Coefficient significance is particularly meaningless, as a *Frequency* coefficient of zero might be a legitimate result in line with the hypothesis if impact changes uniformly. Instead, we transform CARs to absolute values; though we should note again that absolute CARs will be referred to simply as CARs going forward for notational simplicity. To address the transformation, we implement two dummy variables; one to capture intercept differences between negative and positive events, and a second to capture possible slope asymmetries, as explained in section 3.6. Regressions are run using heteroscedasticity-consistent standard errors to account for heteroscedasticity in the sample.[7] Results of the baseline multiple regression follows below:

Table VI
The Baseline Model

Dependent variable		Observations	Model	R^2	$Adj.R^2$	F-statistic
CAR[0,1]		29150	OLS	0.80 %	0.70 %	59.59

Independent Variables	Coefficient	Standard Error	t-statistic	P>|t|	[0.025	0.975]
Intercept	0.0381	0.001	46.100	0.000	0.036	0.040
Frequency	-0.0204	0.003	-7.700	0.000	-0.026	-0.015
Positive	0.0100	0.002	6.144	0.000	0.007	0.013
Frequency x Positive	-0.0152	0.005	-3.116	0.002	-0.025	-0.006

[7] See Appendix C for tests of heteroscedasticity.

Immediately we observe that events of higher levels of announcement frequency *are* associated with lower levels of abnormal returns, under these model specifications. All coefficients are significant at the five percent level despite a poor goodness of fit, with an adjusted R-squared of 0.7%. The intercept dummy, *Positive* is an indicator variable set to one for positive CAR events and zero otherwise. Knowing this we see that, at a frequency of zero, a negative news event is fitted to yield 3.81% absolute abnormal returns, while a positive event yields an abnormal return of 4.81%. As frequency increases, the predicted returns for both classifications move towards zero, though positive news events' fitted returns drop at a rate nearly twice as steep. At a frequency level of 30, a negative news event is projected to yield 3.12% absolute abnormal returns, while a positive event should yield 3.74%. To be clear, *Frequency* and *Frequency x Positive* coefficients represent the change in value per one-hundred announcements. On the surface there seems to exist a clear relationship between *Frequency* and returns, and information flow seems to inversely correlate with impact. Conversely, it is entirely possible, and perhaps plausible, that this relationship is spurious. In the next section we introduce control variables to address the problem of omitted variable bias.

5.2 Expanded Baseline Model with Control Variables

In this section we expand the baseline model from Table VI, by adding control variables. The main control variables include a proxy for company size measured in the market capitalization of each individual firm at the time of the event; a measure for turnover in relative turnover of the security at the time of the event - the relative turnover being turnover over market capitalization; we include indicator variables for industry to capture industry effects; we add indicator variables for time of day effects by classifying whether and event was issued before or after hours, or during the first- or second half of trading; we include indicator variables for each weekday to control for weekday effects; and we include a proxy for company maturity, as measured in total years listed at the time of the event. A table displaying quartile-specific data on all non-indicator control variables is shown below – while tables displaying descriptive statistics for indicator variables are found in appendix B3.

Table VII
Descriptive Statistics - Control Variables

	Frequency	CAR[0,1]	Value	Relative Turnover	Years Listed	Obs.
Panel A: Quartile Means Sorted by Frequency						
Q1	11.13	0.0405	5.94	0.01	7.93	8230
Q2	20.00	0.0378	12.79	0.09	7.88	6729
Q3	27.73	0.0342	17.88	0.08	8.17	7145
Q4	44.49	0.0314	26.02	0.23	8.42	7046
Q1-Q4	-33.35	0.0091***	-20.08***	-0.2232***	-0.4822***	-
Panel B: Quartile Medians Sorted by Frequency						
	Frequency	CAR[0,1]	Value	Relative Turnover	Years Listed	Obs.
Q1	12	0.0214	0.8705	0.0007	6.8137	8230
Q2	20	0.0212	1.6732	0.0014	6.9288	6729
Q3	28	0.0194	2.6909	0.0020	7.2849	7145
Q4	40	0.0179	4.6834	0.0025	8.0329	7046
Q1-Q4	-28	0.0034	-3.8130	-0.0018	-1,2192	-

Table VII reports frequency-sorted quartile mean and median values for *Frequency*, *CAR* and all non-indicator control variables. Quartile one corresponds to the lowest quartile of events sorted by frequency and *CAR*[0,1] is here reported in absolute values to reflect the transformation made in the last section. *Value* is displayed in billions, *Turnover* is displayed in millions, while remaining variables are displayed in normal values. Looking at the *Value* column we immediately notice that larger-valued firms tend to issue more announcements. For both *Turnover* and *Relative Turnover*, we observe a strong increase in trading volume as *Frequency* increases. Finally, *Years Listed* indicates a similar trend, but to a lesser degree. In summary, the table seems to imply that in general, frequent news issuers are more valuable, more mature, and more traded. Most importantly, there is clear change across quartiles – with all quartile differences being significantly different from zero at the one percent level. Overall, we suspect this might indicate that *Frequency* can be explained as a combination of the newly introduced variables.

In the expanded regression we run several different regression models to control for all mentioned effects. All models share the dependent variable *CAR*[0,1] as well as the variables from table VI, i.e., *Frequency*, *Positive* and *Frequency x Positive*. Unique exogenous variables are indicated under 'Control Variables'. Again, regressions are run using heteroscedasticity-consistent standard errors to account for heteroscedasticity in the dataset. Note that to avoid dummy traps, we exclude single dummy-variables from model specifications where exhaustive sets of dummy variables are applied. Specifically, we exclude *Industry 11* (Consumer Discretionary) from *Industries*, we exclude *Second Half* (announcement issued in the second half of trading hours) from the *Time of Day* vector, and we exclude *Monday* from *Weekdays*. In other words, the effect of these variables is absorbed in the intercept of the respective model specifications. Again, *Frequency* coefficients represent a change in frequency per one-hundred announcements. The output of all regression models follow below.

Table VIII
The Baseline Model and Expanded Models

	(1) *CAR*[0,1]	(2) *CAR*[0,1]	(3) *CAR*[0,1]	(4) *CAR*[0,1]	(5) *CAR*[0,1]	(6) *CAR*[0,1]	(8) *CAR*[0,1]
Intercept	0.0381***	0.0352***	0.0382***	0.0381***	0.0431***	0.0252***	0.0262***
	(0.0008)	(0.0010)	(0.0008)	(0.0008)	(0.0010)	(0.0013)	(0.0015)
Frequency	-0.0204***	-0.0218***	-0.0148***	-0.0215***	-0.0191***	-0.0221***	-0.0198***
	(0.0026)	(0.0026)	(0.0027)	(0.0027)	(0.0026)	(0.0027)	(0.0027)
Positive	0.0100***	0.0101***	0.0101***	0.0101***	0.0098***	0.0100***	0.0101***
	(0.0016)	(0.0016)	(0.0016)	(0.0016)	(0.0016)	(0.0016)	(0.0016)
Frequency x	-0.0152***	-0.0153***	-0.0161***	-0.0155***	-0.0148***	-0.0150***	-0.0158***
Positive	(0.0049)	(0.0049)	(0.0049)	(0.0049)	(0.0049)	(0.0049)	(0.0049)
Control Variables							
Time of Day		X					X
Value			X				X
Relative Turnover				X			X
Years Listed					X		X
Industries						X	X
Weekdays							
R-squared	0.80 %	1.00 %	1.50 %	0.90 %	1.10 %	2.00 %	3.20 %
Adj. R-squared	0.70 %	1.00 %	1.40 %	0.90 %	1.10 %	2.00 %	3.10 %
Observations	29 150	29 150	29 150	29 150	29 150	29 150	29 150

There is a lot of information to unpack from table VIII, but the first thing to note is that *Frequency* remains significant across all model specifications.[8] *Value*, the variable we intuitively expected to be most correlated with *Frequency* is

[8] From regressions run with control variables, *Weekdays* (model 7) has the poorest explanatory power and goodness of fit, with all coefficients not significantly different from zero and showing roughly no improvement from the baseline model. This model is therefore excluded from this output and outputs going forward simply because it cannot be fitted within the margins of the document; the outputs for this model are displayed in appendix B2.

significant and seems to decrease the general effect of *Frequency* though there is an enlarged slope asymmetry between positive and negative news indicating that it explains away more of the effect of *Frequency* of negative announcements. *Years Listed* impacts *Frequency* similarly, lowering its modelled effect on abnormal returns but does not lower the asymmetry of positive and negative announcements. *Relative Turnover*, *Time of Day*, and *Industries* also improve goodness of fit - increasingly in that order – and all display significant regression coefficients - though they have negligible effects on *Frequency* as an explanatory variable. For a complete overview of coefficients see appendix B1. To summarize findings, we see that *Frequency* still matters, and is significant at the one percent level in all model specifications – as are remaining baseline variables. For model 8, we find that the *Frequency* coefficient has decreased, while the slope dummy has remained roughly the same. The intercept and dummy intercept have both decreased significantly, but as mentioned previously these are impacted by omitted single variables left out to address the issue of dummy traps, i.e., the intercept values may not be indicative of more profound changes in the model.

6. Robustness Analysis

The coming section addresses points of the analysis we feel are the most prone to have influenced the results of the previous analysis by pure chance. While we would argue all aspects of the study design are defensible, some decisions have been – to different degrees – arbitrary. This section addresses the choices made by exploring viable alternatives.

6.1 Choice of Adjustment Model

As discussed in the section 3.3, the choice of adjustment model for abnormal return computations is not considered to be of strong significance to the outcome of the study; we expected abnormal returns to fall close to actual returns in a two-day study of this sort, and we expected the difference in effect of a factor model versus a mean model to be negligible. In this section we review this decision by investigating the differences across the models. In other words, we compare regressions using the market model, the mean-model, and the factor model implemented using the Fama-French 3-factors. Results follow below.

Table IX
Return-Adjustment Model Comparisons

	Simple *CAR*[0,1]			With Controls *CAR*[0,1]		
	Market	Factor	Mean	Market	Factor	Mean
Intercept	0.0381***	0.0388***	0.0398***	0.0262***	0.0276***	0.0274***
	(0.0008)	(0.0008)	(0.0008)	(0.0015)	(0.0015)	(0.0015)
Frequency	-0.0204***	-0.0213***	-0.0186***	-0.0198***	-0.0202***	-0.0194***
	(0.0026)	(0.0026)	(0.0026)	(0.0027)	(0.0027)	(0.0027)
Positive	0.0100***	0.0092***	0.0091***	0.0101***	0.0094***	0.0094***
	(0.0016)	(0.0016)	(0.0016)	(0.0016)	(0.0016)	(0.0016)
Frequency x Positive	-0.0152***	-0.0147***	-0.0126**	-0.0158***	-0.0152***	-0.0131***
	(0.0049)	(0.0049)	(0.0049)	(0.0049)	(0.0049)	(0.0049)
Control Variables						
Weekdays						
Value				X	X	X
Years listed				X	X	X
Relative Turnover				X	X	X
Industries				X	X	X
Time of day				X	X	X
R-squared	0.80 %	0.70 %	0.60 %	3.20 %	3.10 %	2.90 %
Adjusted R-squared	0.70 %	0.70 %	0.60 %	3.10 %	3.10 %	2.90 %
Observations	29 150	29 150	29 150	29 150	29 150	29 150

The table shows no changes to the significance of the relationship of *Frequency* and abnormal returns, and all variables are deemed significant at the one percent level for both the simple regression model and the expanded model including all relevant control variables. The goodness of fit remains similar across models, as does indicator variables and the intercept. In general, there does not seem to be much difference between the models - certainly not to an extent that would put the results of the prior analysis into question.

6.2 Choice of Time-Period

As discussed in section 3.2, we expect differences in information processing and information availability to affect our data over time. The full sample spans twenty years, and while this period was implicitly chosen when we decided to include the entirety of the *Newsweb* database in our analysis, there may be significant differences within the dataset across time. Without making assumptions about what differences may exist, we conduct testing on the first- and second half of the sample separately to see whether, and how potential breaks in the data across the two halves may have influenced results.

Table X
Time Period Regressions

	Simple *CAR*[0,1]		With Controls *CAR*[0,1]	
	[1999, 2008)	[2008, 2018)	[1999, 2008)	[2008, 2018)
Intercept	0.0345***	0.0391***	0.0242***	0.0300***
	(0.0014)	(0.0010)	(0.0023)	(0.0019)
Frequency	-0.0208***	-0.0173***	-0.0094**	-0.0150***
	(0.0039)	(0.0035)	(0.0042)	(0.0036)
Positive	0.0111***	0.0097***	0.0105***	0.0102***
	(0.0022)	(0.0021)	(0.0022)	(0.0021)
Frequency x Positive	-0.0145**	-0.0163**	-0.0135**	-0.0170***
	(0.0062)	(0.0066)	(0.0061)	(0.0065)
Control Variables				
Weekdays				
Value			X	X
Years listed			X	X
Relative Turnover			X	X
Industries			X	X
Time of day			X	X
R-squared	1.40 %	0.50 %	4.00 %	3.60 %
Adjusted R-squared	1.30 %	0.50 %	3.80 %	3.50 %
Observations	9 507	19 577	9 507	19 577

Again, results are similar across samples, though the 1998 – 2008 sample indicates a better goodness of fit across regressions, and a slightly weaker slope difference between positive and negative events for the simple regression. In general, though, all variables remain significant at the one percent level and there does not seem to be any signs of concerns in the data.

6.3 Frequency Comparison

As explained in the data section, the announcement frequency is measured as the number of announcements over the twelve-month period preceding each event. In other words, it is a dynamic measure and is event specific. The number of months to include in the rolling-estimation window however is completely arbitrary, and again, prudence suggests we validate the choice of a 12-month window by comparing alternatives. Longer cut-offs were never considered, as the rolling window required to compute the measure puts restrictions on where in the sample we can initiate event studies - a larger rolling-windows meaning more data is lost. In the regression below we therefore compare 3-month and 6-month measures of frequency as viable alternatives.

Table XI
Frequency Measure Comparisons

	Simple *CAR*[0,1]			With Controls *CAR*[0,1]		
	3-month	6-month	12-month	3-month	6-month	12-month
Intercept	0.0371***	0.0376***	0.0381***	0.0258***	0.0259***	0.0262***
	(0.0007)	(0.0008)	(0.0008)	(0.0014)	(0.0014)	(0.0015)
Frequency	-0.0679***	-0.0370***	-0.0204***	-0.0622***	-0.0353***	-0.0198***
	(0.0081)	(0.0047)	(0.0026)	(0.0082)	(0.0047)	(0.0027)
Positive	0.0082***	0.0096***	0.0100***	0.0084***	0.0099***	0.0101***
	(0.0013)	(0.0015)	(0.0016)	(0.0013)	(0.0015)	(0.0016)
Frequency x Positive	-0.0343**	-0.0283***	-0.0152***	-0.0377***	-0.0299***	-0.0158***
	(0.0142)	(0.0085)	(0.0049)	(0.0141)	(0.0084)	(0.0049)
Control Variables						
Weekdays						
Value				X	X	X
Years listed				X	X	X
Relative Turnover				X	X	X
Industries				X	X	X
Time of day				X	X	X
R-squared	0.70 %	0.80 %	0.80 %	3.10 %	3.20 %	3.20 %
Adjusted R-squared	0.70 %	0.70 %	0.70 %	3.00 %	3.10 %	3.10 %
Observations	29 150	29 150	29 150	29 150	29 150	29 150

Contrasted with previous robustness checks – though coefficients are similar in terms of significance, as they are mostly significant at the one percent level - there are stark differences in coefficient values. This does not immediately reflect upon the twelve-month measure, as the three alternatives are fundamentally different. To be clear, each alternative operates with a distinct distribution. To illustrate, see figure 3 below.[9]

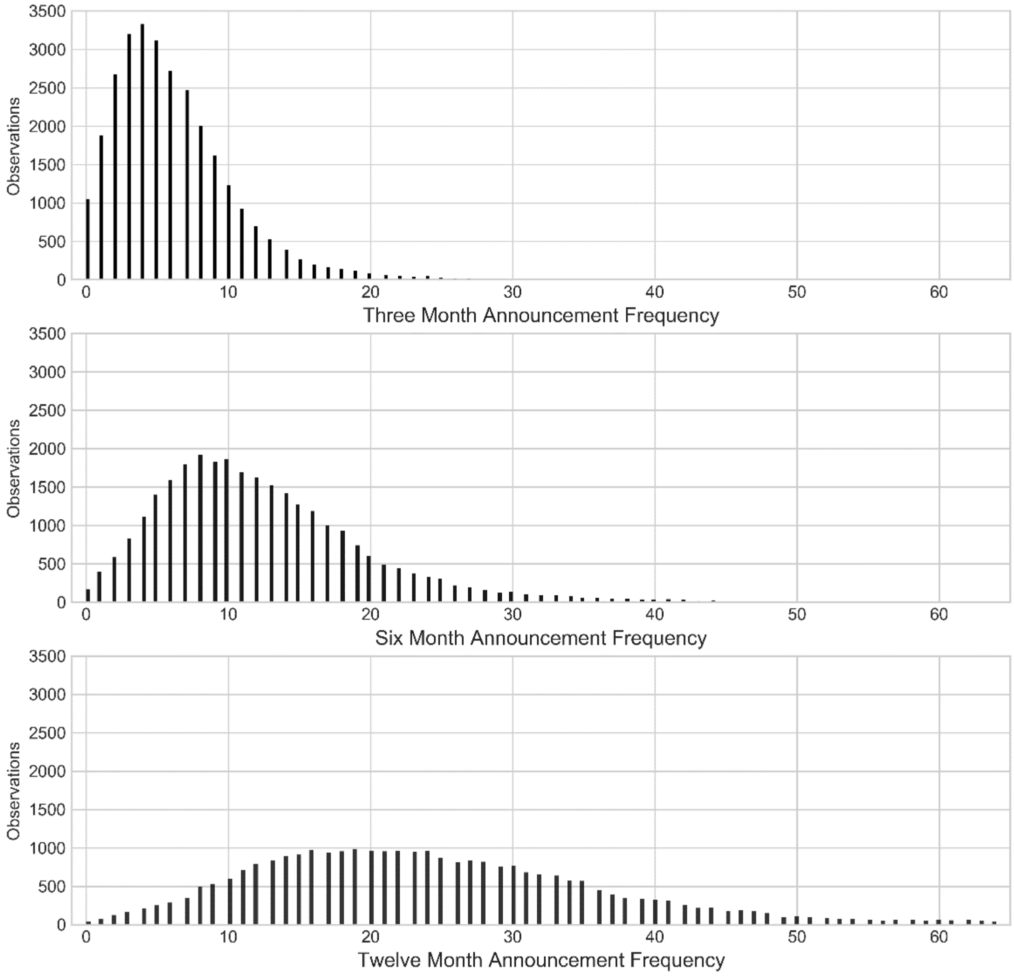


Figure 3: Frequency Measure Distributions

The three-month measure, relative to the twelve-month measure, has no longer-term memory, but seems better at capturing clustering of news announcements. Because of these distinctions they cannot strictly be compared based on model coefficients alone; instead, for purposes of comparison, we apply distribution

[9] For scatter plots of frequency measures in relation to *CAR*[0,1] see appendix A4.

measures of location to their respective models. For fitted model values, see table XII below.

Table XII
Fitted Values – Frequency Measures

Panel A: Distribution Characteristics						
	Mean	Min	25 %	50 %	75 %	Max
12 Month	25	0	16	23	32	151
6 Month	13	0	7	11	16	79
3 Month	6	0	3	5	8	41

Panel B: Fitted Values						
	Positive			Negative		
Frequency	3 Month	6 Month	12 Month	3 Month	6 Month	12 Month
0%	4.5 %	4.7 %	4.8 %	3.7 %	3.8 %	3.8 %
25%	4.2 %	4.2 %	4.2 %	3.5 %	3.5 %	3.5 %
50%	4.0 %	4.0 %	3.9 %	3.4 %	3.3 %	3.4 %
Mean	2.0 %	3.9 %	3.8 %	3.3 %	3.3 %	1.9 %
75%	3.7 %	3.6 %	3.5 %	3.2 %	3.1 %	3.2 %
100%	0.4 %	-0.8 %	-1.2 %	0.8 %	0.6 %	0.8 %

From the fitted table we now see that despite large absolute differences in coefficient values, the fitted values of each respective model are similar. As with previous checks we assume that similarity across specification indicates robustness, and that the slight differences are insufficient to put the results of the main analysis into question.

6.4 Sample Survivorship Bias

As mentioned in section 3.1 our study spans all announcements made by all currently listed companies on the OSE. Currently listed companies is the key phrase for this robustness check, as the specification leaves out a lot of data from all the firms that have been delisted over the years. We think it is fair to assume that this decision introduces some form of survivorship bias to the sample, though exactly how present bias would influence the relationship of *Frequency* and impact is unclear. The reason this data was left out was because – while news and price data for delisted companies is available to us – most of the data is flawed, and much of the control variable data is not available at all. This impacts this robustness test in two ways: first, we have an incomplete sampling of delisted companies meaning inferences draw from this sample may not be representative of the population, and second, we are left with only one vector of controls, namely

Time of Day. The following regression will compare three samples: *live*-, *dead*- and *combined* data, which represent the sample used in the main analysis, the sample of delisted companies, and a combined sample.

Table XIII
Delisted- & Listed-Company Sample Comparisons

	Simple *CAR*[0,1]			With Controls *CAR*[0,1]		
	Dead	Live	Combined	Dead	Live	Combined
Intercept	0.0417***	0.0381***	0.0399***	0.0343***	0.0351***	0.0361***
	(0.0018)	(0.0008)	(0.0008)	(0.0020)	(0.0010)	(0.0009)
Frequency	-0.0017	-0.0205***	-0.0184***	-0.0005	-0.0219***	-0.0195***
	(0.0065)	(0.0026)	(0.0025)	(0.0064)	(0.0026)	(0.0025)
Positive	0.0150***	0.0100***	0.0118***	0.0154***	0.0101***	0.0119***
	(0.0032)	(0.0016)	(0.0015)	(0.0032)	(0.0016)	(0.0015)
Frequency x Positive	-0.0169	-0.0153***	-0.0176***	-0.0174	-0.0154***	-0.0176***
	(0.0116)	(0.0049)	(0.0046)	(0.0116)	(0.0049)	(0.0046)
First Half				0.0063***	0.0018	0.0043***
				(0.0018)	(0.0015)	(0.0012)
Pre Market				0.0087***	0.0039***	0.0047***
				(0.0019)	(0.0014)	(0.0011)
Post Market				0.0505***	-0.0006	0.0154***
				(0.0062)	(0.0023)	(0.0026)
R-squared	0.50 %	0.80 %	0.60 %	1.90 %	1.00 %	0.80 %
Adjusted R-squared	0.40 %	0.70 %	0.60 %	1.80 %	1.00 %	0.90 %
Observations	11847	29153	41000	11847	29153	41000

As shown above, there are fundamental differences between the dead and live sample, and the combined sample seems to mirror the live sample, as it contributes the strong majority of events. *Frequency* is not significantly different from zero for the dead data, though there is a significant *Positive* indicator variable, indicating that positive news tend to be more impactful than negative in this sample. It follows that the dead model has a poorer goodness of fit with an adjusted R-squared of 0.40%. One might conclude from the fact that the combined dataset's *Frequency* coefficient remains significant that the inclusion of dead data would not have impacted the main analysis, but as we mentioned above, a complete sample of dead data would be much larger – and possibly display different characteristics. We therefore recognize that this might have influenced the main analysis, though further analysis is required to ascertain whether this is a sampling phenomenon or not.

6.5 Overlapping Events

The sample which we have used for the analysis throughout the paper consists of roughly thirty-thousand observations, but as explained in section 4.3, it is adjusted to exclude overlapping events. The rationale is to avoid double- or triple-counting returns and overweighting a select number of announcements relative to others. The effect of the adjustment is of course that closely positioned announcements are eliminated - possibly biasing the sample by excluding mostly high-frequency events. In fact, we know it does – in general – as the revised full sample includes almost 20,000 additional observations; has a mean frequency of 29.8 relative to the normal mean frequency of 26.0; and has a median frequency of 25.3 relative to the normal median frequency of 23. To determine what impact this has on model coefficients, we compare the normal sample to the full sample.

Table XIV
Extended Versus Adjusted Sample Comparison

	Baseline *CAR*[0,1]		With Controls *CAR*[0,1]	
	Adjusted	Full Sample	Adjusted	Full Sample
Intercept	0.0381***	0.0391***	0.0262***	0.0279***
	(0.0008)	(0.0006)	(0.0015)	(0.0011)
Frequency	-0.0204***	-0.0225***	-0.0198***	-0.0226***
	(0.0026)	(0.0000)	(0.0027)	(0.0000)
Positive	0.0100***	0.0090***	0.0101***	0.0090***
	(0.0016)	(0.0012)	(0.0016)	(0.0011)
Frequency x Positive	-0.0152***	-0.0133***	-0.0158***	-0.0123***
	(0.0049)	(0.0000)	(0.0049)	(0.0000)
Control Variables				
Weekdays				
Value			X	X
Years listed			X	X
Relative Turnover			X	X
Industries			X	X
Time of day			X	X
R-squared	0.80 %	1.00 %	3.20 %	3.30 %
Adjusted R-squared	0.70 %	1.00 %	3.10 %	3.30 %
Observations	29 150	48 530	29 150	48 530

The two samples exhibit similar characteristics, and both find *Frequency* to be significantly different from zero at the one percent level. For the simple model there are slight differences in slope differentials and the *Positive* intercept, but they really are slight. The problem with using this robustness check for validation of course is that the full dataset is methodically flawed and does not represent legitimate alternative to the design used – this test primarily serves as a sanity check to see that results fall in line with expectations. In conclusion, we do not see any reasons to think that the baseline model was significantly compromised by the slight frequency bias, though it might have had negligible effects.

6.6 Frequency as a Proxy for Information Flow

The last thing we wish to examine is whether *Frequency* is an appropriate measure for information flow by contrasting it to an alternative measure. We have made it clear that our research question focuses on investigating the effects of information flow on abnormal returns, where *Frequency* was selected as the proxy for information flow. It would however, be completely reasonable to use a range of other measures as proxies, and one we had in mind at the start of the process, was to use the number of days between news announcements. This variable, termed *Interval,* would be defined as the number of days between the current event and the last, and in many ways would be a counterpart to *Frequency* in the sense that a large announcement frequency would translate to a smaller mean interval in the sample. Before addressing regression results, we should make clear that contrary to *Frequency* which keeps track of all events in a year*, Intercept* has no long-term memory - and is reset to the value one after every event. To give a sense of data, we have included a plot the two variables relative to announcement cumulative abnormal returns below.

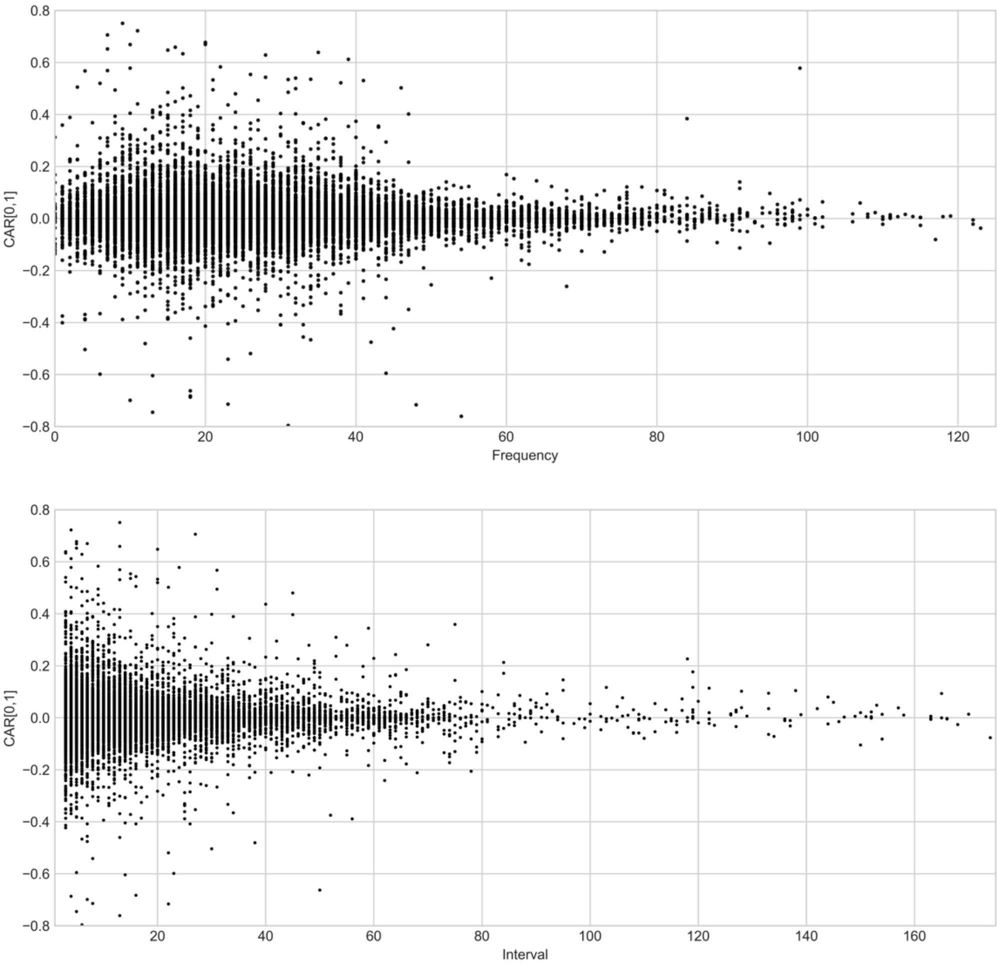


Figure 4: Announcement Returns by Frequency and Interval

The axes of the two figures are different, with the interval axis extends to 175 rather than 125, but the y-axes are identical. As we can tell, *Interval* has a lower median but greater kurtosis than *Frequency*. Full regression outputs for *Interval* are shown below.

Table XV
Interval as Proxy for Information Flow

	(1) *CAR*[0,1]	(2) *CAR*[0,1]	(3) *CAR*[0,1]	(4) *CAR*[0,1]	(5) *CAR*[0,1]	(6) *CAR*[0,1]	(8) *CAR*[0,1]
Intercept	0.0318***	0.0288***	0.0337***	0.0315***	0.0375***	0.0199***	0.0221***
	(0.0005)	(0.0008)	(0.0005)	(0.0005)	(0.0008)	(0.0011)	(0.0014)
Interval	0.0094***	0.0098***	0.0076**	0.0098***	0.0092***	0.0104***	0.0098***
	(0.0030)	(0.0030)	(0.0030)	(0.0030)	(0.0030)	(0.0029)	(0.0030)
Positive	0.0071***	0.0072***	0.0068***	0.0072***	0.0071***	0.0072***	0.0071***
	(0.0008)	(0.0008)	(0.0008)	(0.0008)	(0.0008)	(0.0008)	(0.0008)
Interval x Positive	-0.0083**	-0.0087***	-0.0068**	-0.0086***	-0.0080**	-0.0089***	-0.0083**
	(0.0033)	(0.0033)	(0.0033)	(0.0033)	(0.0033)	(0.0033)	(0.0034)
Control Variables							
Time of Day		X					X
Value			X				X
Relative Turnover				X			X
Years Listed					X		X
Industries						X	X
Weekdays							
R-squared	0.30 %	0.50 %	1.10 %	0.40 %	0.70 %	1.60 %	2.70 %
Adjusted R-squared	0.30 %	0.50 %	1.10 %	0.40 %	0.60 %	1.50 %	2.70 %
Observations	29 150	29 150	29 150	29 150	29 150	29 150	29 150

Table XV shows us that *Interval*, across model specifications, maintains a positive coefficient, significant at the one percent level. This is important to note, because it validates *Frequency* as a measure of information flow to some degree. However, like with the different *Frequency* measures compared in section 6.3, the models are not fully comparable because of the fundamental difference of the measures and it is therefore hard to infer much more without fitting values. Fitted values are shown below:

Table XVI
Fitted Values – Frequency and Interval

Measure		Positive		Negative	
Interval	Frequency	*Frequency*	*Interval*	*Frequency*	*Interval*
3743	0	4.81 %	415.6 %	3.8 %	38.4 %
14	16	4.24 %	5.4 %	3.5 %	3.3 %
13	23	3.99 %	5.3 %	3.3 %	3.3 %
8	25	3.91 %	4.8 %	3.3 %	3.3 %
4	32	3.67 %	4.3 %	3.2 %	3.2 %
3	151	-0.57 %	4.2 %	0.7 %	3.2 %

The table is fitted by two parameters. *Frequency* values are fitted by the leftmost column while *Interval* is fitted by the second to leftmost column. The frequency (interval) level corresponds to the minimum- (maximum-), 25%- (75%-), 50%- (mean-), mean- (50%-), 75%- (25%-), and maximum (minimum) measures of frequency (interval). The table shows some disparity between the two measures for positive news events, while negative news events are close to identical. The first and last rows represent the outliers of the two measures and the focus for interpretation in our view should be on the middle rows. Taking this into account, differences are not substantial, and are in line with each other. In other words, the results indicate that there is a relationship between information flow and abnormal returns on the OSE, where stronger information flow equates to lower average intervals or higher average frequencies of news announcements.

7. Categorical Analysis

One major limitation of the prior analyses is the fact that we look at a broad news sample rather than looking specifically at meaningful subclassifications of news. Ideally, we would have liked to conduct deeper analyses of specific categories and present a more nuanced picture of impact variability and how it changes across categories of announcements. The primary reason we are not able to do this is simply because we do not have rich enough data and can only categorize events by the contents of news headers - which makes manual categorization imprecise and means that inferences drawn could be less meaningful or incorrect. However, we find that there is in fact one category of news that is relatively unambiguous and easy to identify – the category of contract announcements. In this section, as a short tangent to the main analysis, we therefore want to present you with a narrower rendering of the effects of information flow, looking specifically at contracts. In general, we expect there to be close to no noise in the sample of contract announcements, we expect events to be comparable, and we therefore expect to be able to draw inferences which will meaningfully complement or contrast the findings of the main analysis. In table XVII below, we regress a sample of 1615 contract news issuances.

Table XVII

Contract Sample Regressions

	(1) *CAR*[0,1]	(2) *CAR*[0,1]	(3) *CAR*[0,1]	(4) *CAR*[0,1]	(5) *CAR*[0,1]	(6) *CAR*[0,1]	(8) *CAR*[0,1]
Intercept	0.0295***	0.0287***	0.0297***	0.0295***	0.0361***	0.0204***	0.0229***
	(0.0028)	(0.0032)	(0.0029)	(0.0028)	(0.0036)	(0.0040)	(0.0048)
Frequency	-0.0176**	-0.0192**	-0.0082	-0.0178**	-0.0172**	-0.0114	-0.0022
	(0.0082)	(0.0083)	(0.0082)	(0.0082)	(0.0081)	(0.0082)	(0.0088)
Positive	0.0208***	0.0205***	0.0209***	0.0209***	0.0202***	0.0205***	0.0199***
	(0.0050)	(0.0050)	(0.0049)	(0.0050)	(0.0049)	(0.0048)	(0.0048)
Frequency x	-0.0238	-0.0233	-0.0265*	-0.0243	-0.0227	-0.0231	-0.0244*
Positive	(0.0149)	(0.0151)	(0.0147)	(0.0150)	(0.0148)	(0.0144)	(0.0145)
Control Variables							
Time of Day		X					X
Value			X				X
Relative Turnover				X			X
Years Listed					X		X
Industries						X	X
Weekdays							
R-squared	3.00 %	3.30 %	4.10 %	3.10 %	3.60 %	4.40 %	6.60 %
Adj. R-squared	2.80 %	3.00 %	3.90 %	2.80 %	3.40 %	3.80 %	5.60 %
Observations	1 615	1 615	1 615	1 615	1 615	1 615	1 615

Interestingly we see in table XVII that *Frequency* is still an indicator of impact in the baseline model (1), and its sign corresponds with previous analyses throughout the model specifications. However, regression models 3, 6 and 8, including *Value* and *Years Listed,* seems to actually explain away the effects of *Frequency* to the point where the coefficient becomes non-significant. In other words, for this category of news, the inclusion of controls does seem to explain the relationship previously captured by *Frequency*. There are no significant slope differences between positive and negative news for any model specifications, though the *Positive* intercept does remain significant at the one percent level for all models. The latter point however, we would argue, is not surprising, as we would expect contracts to - for the most part – be a non-negative event. In other words, it would be very surprising if *Positive* was *not* significantly different from zero in this sample, as we expect contracts to be an asymmetrically positive category and for negative events to be rare. In summary, looking at a subcategory of news where we know with certainty the observations are all comparable and no large mass of obfuscating noise is present, we find that *Frequency* no longer inversely correlates with abnormal returns. This supports the notion that the relationship found in the main analysis may be a direct result of the composition of news in the sample rather than a general observable relationship.

8. Conclusion

In our analysis of the effects of information flow on abnormal returns, proxied by *Frequency,* as our primary explanatory variable, we find that a relationship seems to exist for the mass of news issued by currently listed companies on the OSE over the last twenty years and that positive events exhibit a stronger relationship in general. This result holds across methodological variations of the *Frequency* measure and for the alternative proxy, *Interval*. The result holds in the first half- and second half of the period independently; they hold for all reviewed return-adjustment models; and they hold when controlling for the frequency bias inferred from the exclusion of - generally higher frequency - overlapping events. Conversely, the inclusion of data from delisted companies seems to reduce the effect of *Frequency*, though the data of delisted companies is incomplete. Including the population sample of news announcements made my delisted companies could feasibly have both strengthened or weakened the relationship, and we therefore argue that the effect of survivorship bias in the sample is inconclusive. The most important, contrasting result to our main analysis, is that the same relationship does not hold for a sub-sample of contract announcements. We interpret this as an indication that information flow's effects on abnormal returns cannot be extrapolated to specific groupings of news within the aggregate sample. Further analysis would be required to fully isolate the differences across categories of news.

9. Shortcomings

While we conclude that a relationship exists between information flow and abnormal event returns, there are several shortcomings to our analysis we would like to address. The first point relates to the fact that the data used for the study is heteroscedastic and non-normal (see appendix C). Heteroscedasticity is acknowledged and responded to with the use of heteroscedasticity robust standard errors in all regressions, but it is less clear how to respond to the issue of non-normality. This is a problem for drawing credible inferences and this should be kept in mind when reading the paper. Second, when looking at reactions to news, we do not expect daily data to be adequate. Having two datapoints represent a forty-eight-hour period following a news-event leaves out a lot of information that could and probably should have impacted our analysis. This ties into a third issue, namely the fact that we look at news events issued by the OSE in isolation. For all we know, a single trading day could contain several news events where only one is issued through *Newsweb* and while there was a clear reaction to the *Newsweb* issue within the first ten minutes, other events may have cancelled out or magnified the effects before trading commencement. The use of daily data also pronounces issues related to return attribution as discussed in section 3.2. When an event is issued one minute after the exchange closes, we choose to attribute the next day's returns as τ_0, while if it was issued ten minutes previously, we would have attributed the whole days returns to the same event. With higher frequency data we would be able to make more nuanced decisions in terms of return attribution and generally would have to make less methodological compromises. Most importantly, while we would defend the design of the study in general, we recognize that slight methodological differences could have had large impacts on general findings.

References

Bagnoli, M., Clement, M. B., and Watts, S. G. (2005). Around-the-Clock Media Coverage and the Timing of Earnings Announcements. McCombs Research Paper Series No. ACC-02-06. Available at SSRN: https://ssrn.com/abstract=570247.

Barber, B. & Lyon, J. D. (1997). Detecting long-run abnormal stock returns: The empirical power and specification of test statistics. The Journal of Financial Economics, 43 (3), 341-372.

Brooks, C. (2014). Introductory Econometrics for Finance. 3rd edition.

Dann, L., Mayers, D., & Raab, R. 1977. Trading rules, large blocks, and the speed of price adjustment. Journal of Financial Economics, 4: 3-22.

DellaVigna, S., and Pollet, J. M. (2009). Investor Inattention and Friday Earnings Announcements. The Journal of Finance, 64 (2), 709-749.

Ederington, L. H., and Lee, J. H. (1993). How Markets Process Information: News Releases and Volatility. The Journal of Finance, 48 (4), 1161-1191.

Fama, E. F. (1970). Efficient Capital Markets: A Review of Theory and Empirical Work. The Journal of Finance, 25 (2), 383-417.

Fama, E. F. (1991). Efficient Capital Markets: II. The Journal of Finance, 46 (5), 1575-1617.

Fama, E. F., Fisher, L., Jensen, C. M., Roll, R. (1969). The Adjustment of Stock Prices to New Information. The International Economic Review, 10 (1), 1-21.

Francis, J., Pagach, D., and Spehan, J. (1992). The Stock-Market Response to Earnings Announcements Release During Trading Versus Nontrading Periods. The Journal of Accounting Research, 30 (2), 165-184.

Hirshleifer, D., Lim, S. S. & Teoh, S. H. (2009). Driven to Distraction: Extraneous Events and Underreaction to Earnings News. The Journal of Finance, 64 (5), 2289-2325.

Hou, K., Peng, L., and Xiong, W. (2008). A Tale of Two Anomalies: The Implications of Investor Attention for Price and Earnings Momentum. Available at SSRN: https://ssrn.com/abstract=976394.

Kadlec, G. B., and McConnell, J. J. (1994). The Effect of Market Segmentation and Illiquidity on Asset Prices: Evidence from Exchange Listings. The Journal of Finance, 49 (2), 611-636.

Kothari, S. P. and Warner, J. B. (2004). The Econometrics of Event Studies. Available at SSRN: https://ssrn.com/abstract=608601 or http://dx.doi.org/10.2139/ssrn.608601

Mackinlay, C. A. (1997). Event Studies in Economics and Finance. The Journal of Economic Literature, 35 (1), 13-39.

McWilliams, A. and Siegel, D. (1997). Event Studies in Management Research: Theoretical and Empirical Issues. The Academy of Management Journal, Vol. 40, No. 3. 626-657.

Marshall, Ben R. and Nguyen, Nhut H. and Visaltanachoti, Nuttawat, A Note on Intraday Event Studies (August 8, 2017). Available at SSRN: https://ssrn.com/abstract=3015618 or http://dx.doi.org/10.2139/ssrn.3015618.

Mitchell, M., & Netter, J. M. 1989. Triggering the 1987 stock market crash: Antitakeover provi- sions in proposed House Ways and Means Committee tax bill. Journal of Financial Eco- nomics, 24: 37-49.

Peterson, P. P. (1989). Event Studies: A Review of Issues and Methodology. Quarterly Journal of Business and Economics, 28 (3), 36-66.

Ryngaert, M., and Netter, J. 1990. Shareholder wealth effects of the 1986 Ohio antitakeover law revisited: Its real effects. Journal of Law, Economics, & Organization, 6: 253-262.

Zhang, X. F. (2006). Information Uncertainty and Stock Returns. The Journal of Finance, 61 (1), 105-137.

Ødegaard, B. A. (2018). Fama-French Factors. Available at: http://finance.bi.no/~bernt/financial_data/ose_asset_pricing_data/index.html. Accessed 07/02/2018.

Appendix A

A1. Noise List

FINANCIAL CALENDAR
INVITATION TO PRESENTATION
ORDINARY GENERAL MEETING
PRESENTATION OF RESULTS
INVITATION TO 1Q
INVITATION TO 2Q
INVITATION TO 3Q
INVITATION TO 4Q
ANNUAL GENERAL MEETING
INVITATION TO EARNINGS RELEASE
INVITASJON TIL PRESENTASJON
FINANSIELL KALENDER
ORDINÆR GENERALFORSAMLING
PRESENTASJON AV RESULTAT
RENTEREGULERING
NY RENTE
EX-DIVIDEND

To address the issue of overlapping events, we implemented a two-stage solution as discussed in section 4.3. The first part included removing any non-relevant news issuances – noise - which would almost certainly not include any new information and therefore should not earn any abnormal returns. In practice, news which headers included any of the phrases above were removed.

A2. Simple T-tests

T-tests were performed on the aggregate mean measures of the $CAR[0,1]$-, AR_0-, and AR_1 variables.

	Mean	t-stat	5% Critical Value	1% Critical Value
$CAR[0,1]$	0.004431	9.42651	1.960061	2.57603
AR_0	0.004367	12.175815	1.960061	2.57603
AR_1	0.000064	0.210329	1.960061	2.57603

As shown in the table, AR_1 was found to not be significantly different from zero, while the remaining two were found to be significantly different from zero at the one percent level.

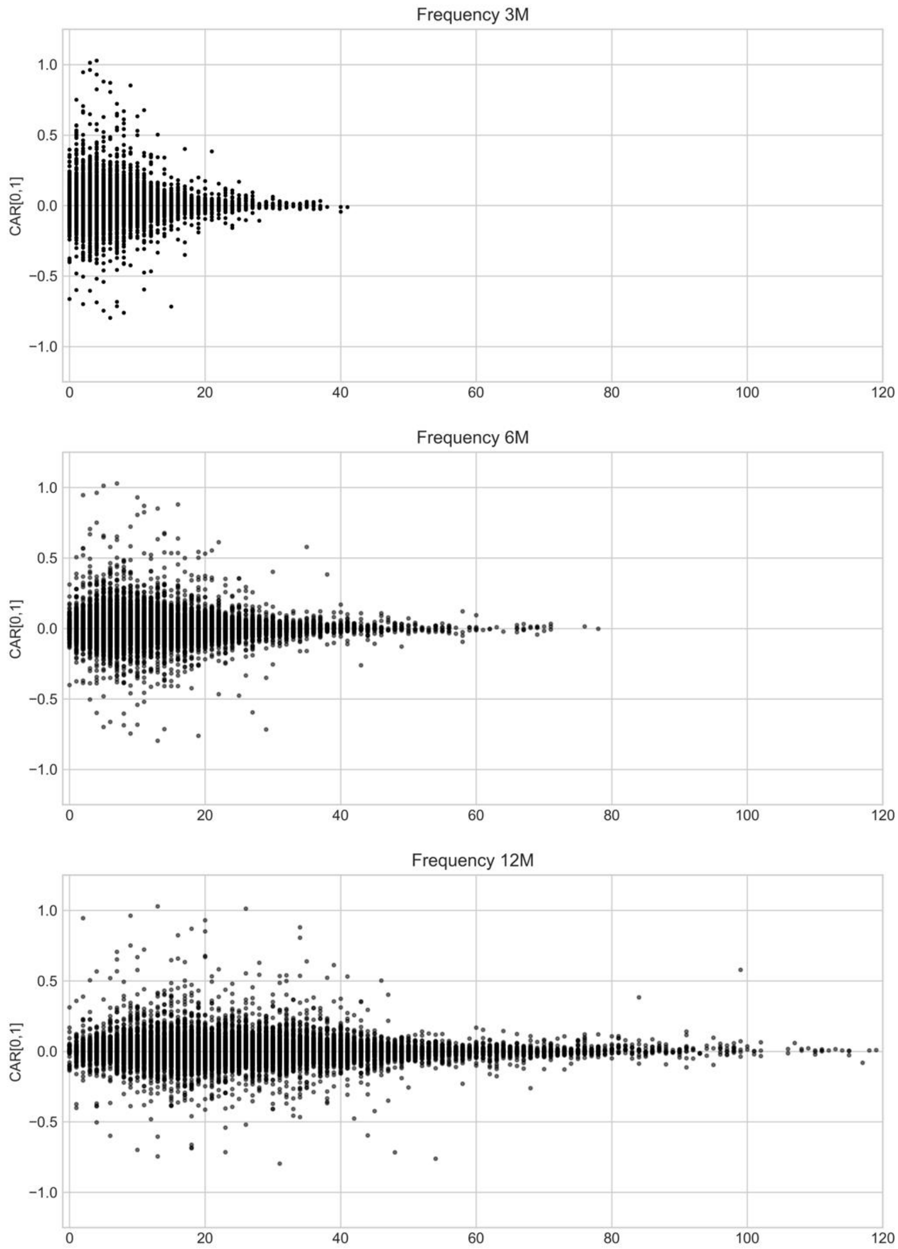
A3. F-tests of Quartile Differences in Variance

A comparison of frequency-quartile variance levels found that the first quartile is shown to be significantly different from all other quartiles of higher announcement frequencies:

	Q1	Q2	Q3	Q4
Q1	1			
Q2	1.54***	1		
Q3	2.03***	1.32***	1	
Q4	2.25***	1.46***	1.11***	1

This result holds across quartiles and is interpreted as an indication that there may be a relationship between *Frequency* and impact. This however, is not included as proof of any hypothesis – it is only meant to serves as a preliminary sanity test to show that there might be something worth investigating in the general sample.

A4. Alternative Frequency Measures



The plot shows cumulative abnormal returns fitted against *Frequency* for three different methodological variations of the variable. The first uses a three month-, the second uses a six month-, and the last uses a twelve-month rolling-window to estimate the announcement rate of a firm at the time of an individual event.

Appendix B

B1. Full Baseline Regression Output

The Baseline- and Expanded Models

	(1) *CAR*[0,1]	(2) *CAR*[0,1]	(3) *CAR*[0,1]	(4) *CAR*[0,1]	(5) *CAR*[0,1]	(6) *CAR*[0,1]	(8) *CAR*[0,1]
Intercept	0.0381***	0.0352***	0.0382***	0.0381***	0.0431***	0.0252***	0.0262***
	(0.0008)	(0.0010)	(0.0008)	(0.0008)	(0.0010)	(0.0013)	(0.0015)
Frequency	-0.0204***	-0.0218***	-0.0148***	-0.0215***	-0.0191***	-0.0221***	-0.0198***
	(0.0026)	(0.0026)	(0.0027)	(0.0027)	(0.0026)	(0.0027)	(0.0027)
Positive	0.0100***	0.0101***	0.0101***	0.0101***	0.0098***	0.0100***	0.0101***
	(0.0016)	(0.0016)	(0.0016)	(0.0016)	(0.0016)	(0.0016)	(0.0016)
Frequency x	-0.0152***	-0.0153***	-0.0161***	-0.0155***	-0.0148***	-0.0150***	-0.0158***
Positive	(0.0049)	(0.0049)	(0.0049)	(0.0049)	(0.0049)	(0.0049)	(0.0049)
First Half		0.0022**					0.0016*
		(0.0010)					(0.0010)
Pre Market		0.0065***					0.0063***
		(0.0009)					(0.0009)
Post Market		-0.0007					-0.0006
		(0.0012)					(0.0012)
Relative Turnover				0.0022***			0.0019**
				(0.0008)			(0.0008)
Value			-0.0001***				-0.0001***
			(0.0000)				(0.0000)
Years Listed					-0.0007***		-0.0004***
					(0.0001)		(0.0001)
I1						0.0006	0.0129***
						(0.0015)	(0.0016)
I10						0.0219***	0.0204***
						(0.0014)	(0.0014)
I2						0.0101***	0.0125***
						(0.0018)	(0.0018)
I3						0.0183***	0.0175***
						(0.0018)	(0.0018)
I4						-0.0050***	-0.0050***
						(0.0019)	(0.0019)
I5						0.0071***	0.0064***
						(0.0022)	(0.0022)
I6						0.0182***	0.0181***
						(0.0013)	(0.0014)
I7						0.0007	0.0019
						(0.0012)	(0.0012)
I8						0.0076***	0.0085***
						(0.0016)	(0.0016)
I9						0.0103***	0.0102***
						(0.0012)	(0.0012)
R-squared	0.80 %	1.00 %	1.50 %	0.90 %	1.10 %	2.00 %	3.20 %
Adj. R-squared	0.70 %	1.00 %	1.40 %	0.90 %	1.10 %	2.00 %	3.10 %
Observations	29150	29150	29150	29150	29150	29150	29150

The regression output above shows all regression models from the main analysis – including coefficient values and significance levels.

B2. Weekday Regression

The Baseline- and Weekday Model

	(1) *CAR*[0,1]	(7) *CAR*[0,1]
Intercept	0.0381***	0.0382***
	(0.0008)	(0.0012)
Frequency	-0.0204***	-0.0204***
	(0.0026)	(0.0026)
Positive	0.0100***	0.0100***
	(0.0016)	(0.0016)
Frequency x Positive	-0.0152***	-0.0153***
	(0.0049)	(0.0049)
Friday		-0.0003
		(0.0012)
Thursday		0.0006
		(0.0012)
Tuesday		-0.0004
		(0.0012)
Wednesday		-0.0004
		(0.0013)
R-squared	0.80 %	0.80 %
Adj. R-squared	0.70 %	0.70 %
Observations	29150	29150

Because of margin specifications, we were not able to fit all regression models in a single table. We therefore decided to leave out the model with the lowest improvement from the baseline model; the *Weekdays* regression model. This regression model is instead shown here.

B3. Indicator Variable Quartile Means

Quartile Means Sorted by Frequency - Time of Day

	Frequency	Pre Market	Post Market	First Half	Second Half	Obs
Q1	11,13	37,7 %	10,7 %	24,0 %	27,5 %	8 230
Q2	20,00	41,3 %	8,4 %	24,3 %	25,9 %	6 729
Q3	27,73	43,3 %	7,7 %	25,1 %	23,8 %	7 145
Q4	44,49	45,6 %	6,9 %	24,9 %	22,5 %	7 046
Q1-Q4	-33.35	-7.9%***	3.8***	-0.9%	5.0%***	7 046

Quartile Means Sorted by Frequency - Industries

	Frequency	I1	I2	I3	I4	I5	I6
Q1	11,13	0,3 %	4,1 %	6,3 %	2,7 %	4,0 %	28,3 %
Q2	20,00	0,3 %	4,7 %	4,5 %	1,0 %	2,5 %	34,7 %
Q3	27,73	1,2 %	4,6 %	3,9 %	0,2 %	2,4 %	32,1 %
Q4	44,49	4,4 %	6,6 %	1,9 %	0,0 %	2,6 %	31,8 %
Q1-Q4	-33.35	-4.1%***	-2.4%***	4.4%***	2.8%***	1.4%***	-3.5%***
	Frequency	I7	I8	I9	I10	I11	Obs
Q1	11,13	7,4 %	7,0 %	22,7 %	12,5 %	4,6 %	8230
Q2	20,00	4,4 %	8,8 %	21,5 %	13,9 %	3,7 %	6729
Q3	27,73	3,6 %	10,3 %	22,8 %	15,5 %	3,4 %	7145
Q4	44,49	5,4 %	7,8 %	22,0 %	16,2 %	1,5 %	7046
Q1-Q4	-33.35	2.0%***	-0.8%*	0.7%	-3.6%***	3.1%***	7046

Quartile Means Sorted by Frequency - Weekday

	Frequency	Monday	Tuesday	Wednesday	Thursday	Friday	Obs
Q1	11,13	16,7 %	18,7 %	20,9 %	22,5 %	21,3 %	8230
Q2	20,00	17,5 %	19,5 %	20,3 %	22,1 %	20,7 %	6729
Q3	27,73	17,5 %	19,7 %	21,0 %	20,8 %	21,0 %	7145
Q4	44,49	18,3 %	19,1 %	21,4 %	21,4 %	19,7 %	7046
Q1-Q4	-33.35	-1.6%***	-0.5%	-0.6%	1.1%	1.6%**	7046

B4. Price Data

Date	AFG	AFK	…	WWI-OS	WWIB-OS
26/01/1998	5.73	506.97	…	70	70
27/01/1998	5.73	506.97	…	69	70
28/01/1998	5.73	506.97	…	68.75	70
29/01/1998	5.73	506.97	…	69	70
…	…	…	…	…	…
23/01/2018	129.5	3380	…	262.5	260
24/01/2018	132	3380	…	263.5	259
25/01/2018	131	3380	…	259.5	253
26/01/2018	118.5	3340	…	259	254

The above table shows an example of the price data we use for event study procedures. The prices displayed are daily closing prices.

B5. News Data

Statoil	Statoil	Statoil	Statoil
10/06-2002	10/06-2002	03/06-2002	03/06-2002
08:30:32	08:30:03	09:12:09	09:03:45
STL	STL	STL	STL
STL - UK GAS CONTRACT FOR STATOIL (OBI)	STL - STATOIL MED STORT GASSALG TIL UK (OBI)	STL - STATOIL GETS GO-AHEAD FOR SNØHVIT (OBI)	STL - STATOIL FÅR GRØNT LYS FOR SNØHVIT (OBI)

News data contains the company name, the date of the announcement, the time of the announcement, the company ticker, and the title of the announcement. The complete dataset contains all news related to a single company and includes a separate .csv file for each relevant company.

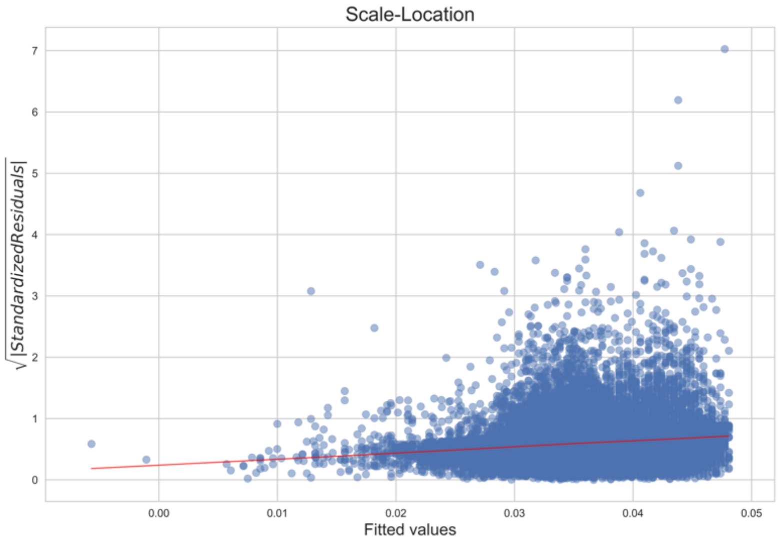
B6: Pricing Factors Daily (Including Fama-French Factors)

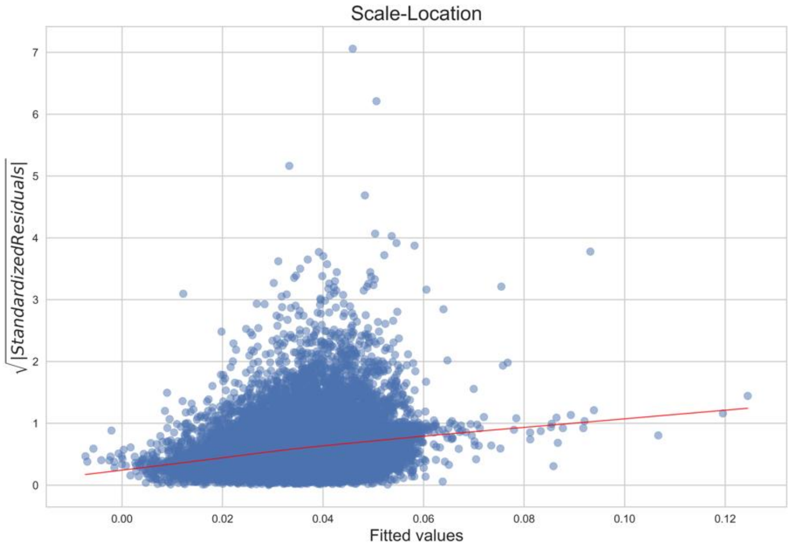
Date	SMB	HML	PR1YR	UMD	Rf(1d)	EW	VW	Allshare
19861201	-0.004	0.000	-0.001	0.000	0.001	-0.003	-0.004	-0.003
19861202	-0.006	-0.002	0.013	0.006	0.001	-0.015	-0.014	-0.012
19861203	-0.003	-0.004	0.011	-0.002	0.001	-0.024	-0.019	-0.019
19861204	-0.007	0.000	0.015	0.013	0.001	-0.009	-0.006	-0.002
19861205	0.000	-0.005	-0.011	-0.004	0.001	0.006	0.002	0.000
19861208	0.005	-0.001	-0.004	0.003	0.001	-0.009	-0.010	-0.009
19861209	0.010	0.008	0.001	0.003	0.001	-0.009	-0.013	-0.017
19861210	0.000	0.005	-0.002	0.009	0.001	0.006	0.007	0.008
19861211	-0.017	-0.015	-0.010	-0.017	0.001	0.004	0.010	0.012

The Fama-French 3-factor portfolios are calculated by Bernt Arne Ødegaard.

Appendix C

Before starting our analysis, we conducted diagnostics testing on our data samples. The first regression assumption, $E(\varepsilon_i) = 0$ is met from the fact that we include an intercept in all regression models (Brooks 2014). The second assumption, $\sigma_u^2 < \infty$, that the variance of the residuals is finite and constant, i.e., that the residuals are homoscedastic is not met as shown in the plots below. The first plot shows fitted residual values of our baseline regression, while the second plot shows fitted values of the extended regression (model 8) including controls.



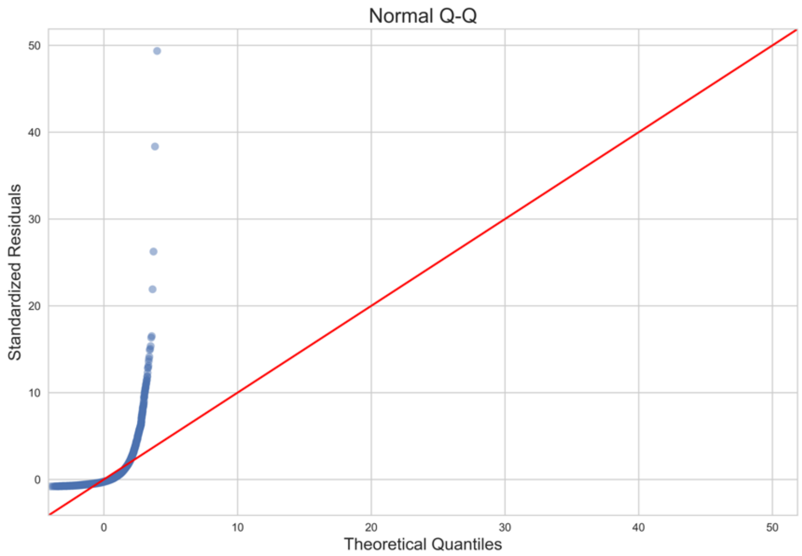


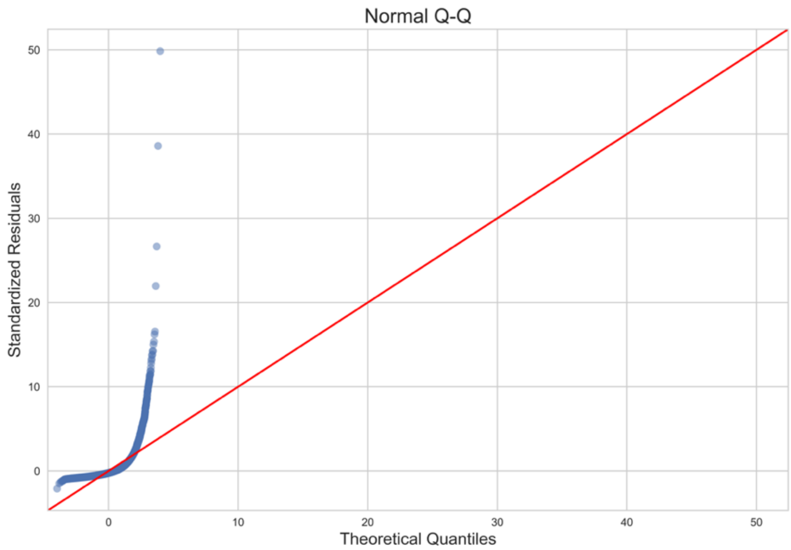
We observe residuals spread unevenly across fitted values, growing larger for higher values of predictors in the model. In other words, the plots visually seem to indicate the presence of heteroscedasticity which is confirmed by running a Breusch-Pagan test for heteroscedasticity as shown below.

Breusch-Pagan test for Heteroscedasticity

	Baseline Model	Model with Controls
Langrange Multiplier Statistic	80.2352	78.1661
Lagrange Multiplier P-Value	0.0000	0.0000
F-Statistic	40.2256	39.1855
F-Statistic P-value	0.0000	0.0000

To address the issue of heteroscedasticity in our data, we run all regressions using heteroscedasticity-robust standard errors. While the third and fourth assumptions of the CLRM model do not affect us, the fifth assumption does, i.e., the assumption of normally distributed residuals. To test, we first computed Q-Q plots of the dataset for the baseline- and extended model. Plots are shown below.



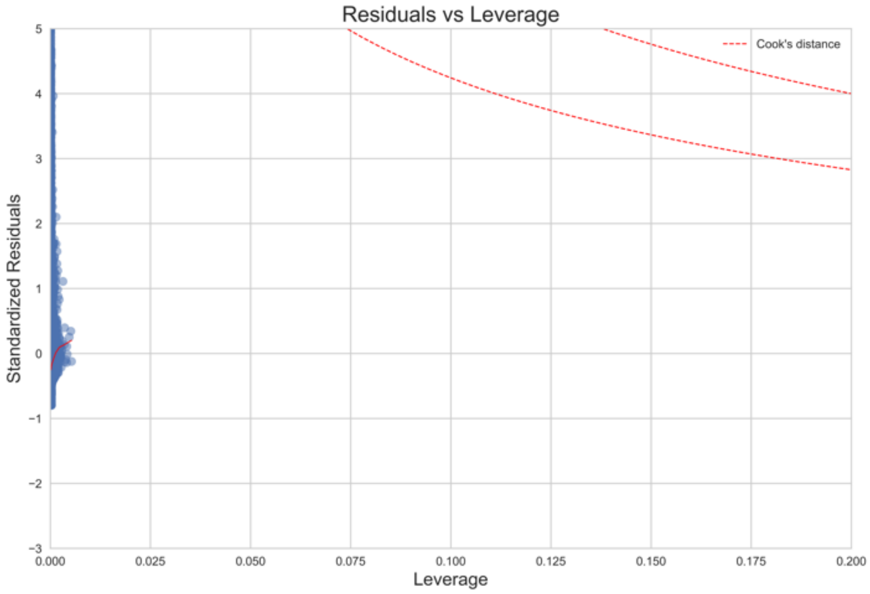


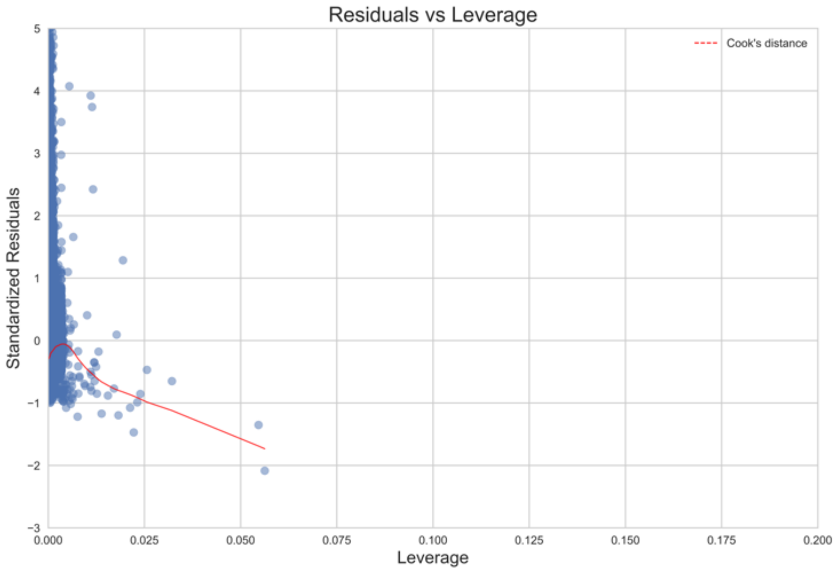
The plots do not correspond with our expectations of a normal dataset as the residuals do not hug the line. We therefore conduct standardized testing for normality using the Jarque-Bera test.

Jarque-Bera Statistics from Regression Models

	test-statistic	p-value
Model 1	139803947.8	0.0000
Model 2	…	0.0000
Model 3	…	0.0000
Model 4	…	0.0000
Model 5	…	0.0000
Model 6	…	0.0000
Model 7	…	0.0000
Model 8	149006540.5	0.0000

We know that the critical value of the Jarque-Bera test is 4, where the null hypothesis is normality. This hypothesis, in other words, is firmly disproven. Lastly, for diagnostics purposes, we plotted standardized residuals over leverage. These plots are helpful to examine the influence of outliers on the model fit. Figures are shown below for the baseline and extended model.





Although we observe extreme values, there seem to be no influential outliers in the sample that alter the regression model to any significant extent, as none of the extreme values exceed Cook's distance.