ERASMUS UNIVERSITY ROTTERDAM
Erasmus School of Economics
Bachelor Thesis: International Bachelor Economics and Business Economics

Media sentiment and the IPO Quiet Period Anomaly

Student Name: Adwaith Nair
Student ID number: 561959

Supervisor: dr. R. de Bliek
Second Assessor: dr. F. Core

Date Final Version: 04/07/2023

ABSTRACT

This research investigates the quiet period anomaly - a phenomena where IPOs tend to have abnormal returns after the 40 day SEC regulatory period. It examines the relationship between Twitter sentiment - calculated using Roberta, an NLP software - and abnormal returns at the end of the quiet period by using multivariate regression models with industry, audit company and a binary variable indicating whether a company is classified as an EGC as the control variables. The results indicate that a higher neutral and negative sentiment is significantly associated with a decrease in abnormal returns while positive sentiment does not have a significant relationship. The inclusion of control variables strengthen this effect. The volume of tweets, however, do not seem to have an effect on abnormal returns. Additional sentiment calculations were conducted using VADER - a bag-of-words approach - which yielded conflicting results, stating that an increase in positive sentiment decreases CMAR and an increase in negative sentiment increases CMAR.

Table Of Contents

I. INTRODUCTION

Does media sentiment of a company shape stock market returns? It has been long considered that a key cause of volatility in the stock market, and any investment, is relevant media exposure. From financials to corporate strategy, to management and legal changes, press releases have been documented to generate higher levels of valuation uncertainty (Neuhierl et al., 2013). In a market where uncertainty is already high, IPOs in the United States are imposed by the Securities Exchange Commission (SEC) to abide to a "quiet period" - a period of 25-day, post-IPO, where firms and their underwriters are banned from publishing information that is not included in the prospectus. The effectiveness of this rule, however, has been disputed due to the tendency for significant abnormal returns towards the end of the period (Bradley et al., 2003). Thus, to better understand this puzzle and shed light on the IPO cycle, this paper analyzes factors – in particular media tone (sentiment) – that may effect post quiet period returns.

Standard models in economics state that markets are efficient (Fama, 2013) and that asset prices reflect all available information. However, the observation of market anomalies are inconsistent with the Efficient Market Hypothesis (EMH) even in IPOs due to: underpricing, long term performance and the quiet period puzzle. Abnormal IPO returns have been extensively studied with papers explaining factors such as underwriter rank (Carter, 1998) and social media sentiment (Liew & Wang, 2016) that affect underpricing and performance. Media sentiment has also been previously researched with reference to underpricing, with Bajo and Raimondo (2017) concluding that positive media sentiment is associated with a greater extent of underpricing. When it comes to the quiet period puzzle, however, academic discourse is limited. The first paper to mention the quiet period and the accompanying puzzle was Bradley, Jordan & Ritter (2003) who analysed the expiration of the quiet period and concluded that the 76% of firms that initiated analyst coverage immediately experienced a five-day abnormal return of 4.1%. Attempts to explain this include Cedergren (2014) who found that underwriters are still participating through media during the IPO quiet period despite regulatory restrictions. With regards to the media, Bushee et al. (2020) finds that media coverage during the quiet period is associated with more retail investor purchases, but no conclusion has been made regarding its effect on the quiet period puzzle.

Existing papers, which present an association between the media and IPO returns, seem to suggest that the way in which sentiment on a company is presented by the media may help shape investor beliefs and in turn drive post quiet period returns. At the time of writing, however, this association is yet to be studied empirically, leading to the research question: How does media sentiment affect abnormal returns during the quiet period?

The relevant U.S. IPO data for this research, including the relevant company information and post quiet period stock return data is obtained from Wharton Research Data Service (WRDS) database. Market-adjusted returns (CMAR) are calculated and used as a measure of abnormal returns. With regards to media sentiment, tweets for sentiment analysis are scraped using snscrape. Sentiment analysis is conducted using natural language processing – a process through which computers manipulate language text to conduct analysis (Chowdhury, 2003.) - on python using a BERT transformer. A regression model is finally utilized to test for an association between the sentiment score and CMAR.

In summary we find evidence that media sentiment during the quiet period has a significant impact on post-IPO performance. Specifically, firms with higher neutral and negative sentiment scores are significantly associated with a decrease in abnormal returns, when other factors (such as industry, auditor rank etc.) are controlled for. Positive sentiment scores, however, do not have significant impact. The volume of tweets are also found to not have an impact. The results of this research can help better inform market participants to make accurate decisions.

In the following sections, a brief background into the topic and existing research are first provided. The sources of data and methodology are then described, before describing the results for the sample of IPOs. Then, some methodological variations are introduced before finally discussing the results and contributions of this research.

II. Theoretical framework

II.A. The Efficient Market Hypothesis and IPO anomalies

Often considered as a cornerstone of modern finance, the Efficient Market Hypothesis (EMH) states that markets are fully efficient as "prices fully reflect all available information" (Fama, 1970, p. 383). Through empirical tests on the weak form, semi-strong form and strong form, Fama concluded that abnormal returns are impossible for stocks. This, however, is largely disputed with the existence of market anomalies - situations where stocks tend to deviate from the EMH. A notable example of an anomaly in the financial market is the January effect, where Keims (1983) conducted a regression analysis between abnormal returns and the market value of NYSE and AMEX common stocks to establish that small company stocks generate more return than other classes in the first weeks of January. Theories to explain these anomalies tend to stem from behavioral psychology including, for instance: the overreaction hypothesis, which states that abnormal returns for stocks are caused by individuals overreacting to unexpected and dramatic news (DeBondt and Thaler, 1985).

With regards to Initial Public Offerings (IPOs), Fama and EMH proponents would argue that once a company is traded, the stock price should reflect its intrinsic value and capture all available information. However, in a market where uncertainty is already high, IPOs are also subject to unexplained anomalies that result in abnormal returns, much of which is explored in the seminal work of Ritter and Welch (2002). A plethora of research has also attempted to explain factors contributing to these anomalies.

For example, a phenomena that has been under scrutiny to extensive research is short run underpricing - where companies that have recently undergone an IPO, experience significant price increases on the first day of trading (Ritter, 1984). In a sample of IPOs from 1980 to 2001, it was documented that 70% of the IPOs end the first day at a closing price greater than their opening price. Forces that influence this are captured in the review of Katti and Phani (2016) and include issue, firm and economic factors such as: issue size, underwriter rank and industry cycle.

Another significant IPO anomaly is long run underperformance which, as documented by Loughran & Ritter (1995), show that IPOs on average tend to underperform the market significantly for the five years after issuance. Similar factors were seen to contribute to this, with Carter, Dark and Singh (2002) establishing that the underperformance of IPO is less severe for IPOs for whom underwriter ranks are higher and Ang and Boyer (2009) concluding - through a longitudinal study - that there exist industry differences with IPOs in new industries outperforming those in established industries.

II.B. The Quiet Period Anomaly and Policy Background

The existence of these anomalies has led to growing research and interest on IPOs in academic literature. However, an anomaly which is studied less extensively is the Quiet Period Puzzle - marked by abnormal returns observed at the end of a regulatory period in the United States.

The quiet period is a regulatory restriction imposed by the Securities and Exchange Commission (SEC) on companies that take part in an IPO. It is a period of time where IPOs are considered to still be in registration and where affiliated analysts are prohibited from issuing recommendations and from publicly making forward looking statements or opinions on the company (Bradley, Jordan & Ritter, 2003). Coverage, that is often favorable to a company, is typically initiated by insiders and underwriters immediately at the end of the period.

Through this regulation, the SEC attempts to ensure that all investors have access to the same information. Thus, in an efficient market, no abnormal returns are expected at the end of the quiet period as all available information should be captured by the prices. However, as documented by Bradley, Jordan & Ritter (2003), this is not the case. In their paper, the quiet period for IPOs from 1996 to 2000 is examined through an event study, and it is concluded that a 5-day abnormal return of 4.1% is concentrated just before the expiration of the quiet period. This is in line with other findings such as Highfield, Lach and White (2008) who found that for a 2-day window consisting the day before and the day of the quiet period expiration, a positive cumulative market-adjusted return can be observed. While analysing factors that contribute to this, they further establish, through multivariate regression analysis, that IPO size and the 'number of buy ratings' significantly affect the quiet period IPO return.

To improve the effectiveness of the quiet period, the relevant regulations have been under constant scrutiny by the SEC, and have thus changed over time. The most notable recent changes are: amendments in 2002 which increased the duration of the quiet period from 25 to 40 days; changes in 2006 which allowed for better communication to reach investors before an IPO; and the JOBS act of 2012 which created a category of Emerging Growth Companies (EGCs) that are allowed to communicate with institutional buyers during the quiet period (Thompson, 2017).

To test the robustness of the quiet period anomaly to regulatory changes, Bradley, Jordan, Ritter & Welch (2004) conducted follow up research on their seminal paper. Using a similar methodology to test a sample of companies (from 2001 to 2002) who experienced an IPO in the extended period of 40 days, they concluded that there still is abnormal returns for firms with coverage, albeit at a smaller percentage. The persistence of these returns despite regulatory shifts, suggest that unregulated may factors play a role.

As such, this paper specifically focuses on the role of media, a source which is not restricted during the quiet period and attempts to find whether an association exists between the media and the reported abnormal returns.

II.C. The Role of Media

The role of the media has become increasingly relevant in the past decades. Media informs and entertains the audience and through this exchange of information plays the important role of shaping public beliefs and attitudes. In particular, the growth of social media - defined as internet based channels of communication facilitation interactions and information transfer among individuals (Carr & Hayes, 2015) - has been under scrutiny to research. A plethora of literature has thus researched the effect of media, reporting that apart from enhancing accessibility of information, it can also lead to cognitive bias and influence the behavior of the audience.

This influence on behavior has been studied in several strands of literature. In social psychology, Greitemeyer (2011) reviews the role of media violence and through the General Learning Model shows that exposure to prosocial content in media - especially through music and video games - increases prosocial empathy and helping behavior amongst indivduals. In the context of politics, DellaVigna and Kaplan (2007) used a natural experiment to study the effect of media bias on voting behavior in the United states between 1996 and 2000. From a database of voting data in 9256 towns, their findings imply that the introduction of Fox News - which is right leaning - resulted in a significant increase on the Republican vote share by 0.4-0.7%.

This effect on behavior is also captured within Finance. Citing time and bounded rationality as barriers which prevent humans from processing a plethora of information, Barber and Odean (2008) argue that retail investors are net buyers of attention-grabbing stocks. By first sorting investments on whether relevant stocks were in the daily news feed and then calculating buy-sell imbalances for each firm's stock, the authors conclude that investors are significantly more likely to be net buyers of stocks that are in the media than those that are not.

Shen, Urquhart and Wang (2019) further establishes this effect as a result of social media by researching the link between investor attention and Bitcoin returns. With the support of granger causality tests, they conclude that number of tweets is a significant driver of next day trading volume in Bitcoin. This is in line with research on IPOs, where Kwan (2015) analyzed over 400 million tweets on companies that had an IPO in 2009 and concluded that an increase in the number of tweets correlates with higher first day returns after IPO. In view of the this, it is hypothesized that exposure to tweets leads to an increase in trading volume and abnormal returns at the end of the quiet period:

H1: An increase in the quantity of tweets on a firm during the quiet period leads to an increase in abnormal returns at the end of the period.

II.D. Media Sentiment and the Quiet Period Anomaly

A stream of literature has taken the role of media further by arguing that it is not just media volume but also media sentiment which affects individual behavior.

The forerunners to literature on media sentiment can be traced back to the field of linguistics where Wiebe (1994) discussed subjectivity in private states of characters in prose. Dave et al (2003) later coined the phrase 'opinion mining', defined as a process that identifies the distinction between positive and negative product reviews. This eventually led to the modern term sentiment analysis, defined as the computational study of opinions, attitudes, and emotions, expressed in a text (Medhat, Hassan & Korashy, 2014), from which we derive a definition for media sentiment; the analysis of opinions in media. Today, analysis of sentiment has quickly become one of the fastest growing research areas in computer science. It is facilitated by the use of Natural language Processing (NLP) strategies, like the Bag-Of-Words approach, and deep language models, like BERT and RoBERTa (Liu et al 2019).

Within finance literature, sentiment analysis has become particularly important while analysing social media, in particular twitter sentiment. Ranco, Aleksovski, Caldarelli, Grcar and Mozetic (2015) conducts an event study relating twitter sentiment and stock returns. By looking at a period of 15 months and looking into 30 companies, they conclude that sentiment of tweets (through NLP) at the period of peak twitter volume is significantly associated with the direction of cumulative abnormal returns.

This is in line with the analysis of Media in the context of IPO literature. Bajo and Raimondo (2017) conduct an analysis of media sentiment and IPO underpricing on 2800 US IPOs and over 27,000 newspaper articles. Through textual analysis - using a bag of words approach - and regression, it is concluded that positive sentiment is positively associated with IPO underpricing and that this effect is stronger when news is reported close to the IPO date.

Liew and Wang (2016), who conducts a similar analysis focusing on social media sentiment and IPO returns and find a significant positive relationship between prior days' tweet sentiment and next-day IPO returns. Bushee, Cedergren & Michels (2019), further focuses on the quiet period and through multiple regression models find that more media coverage during the period is associated with more purchases by retail investors during the period. However, they make no conclusion regarding the abnormal returns at the time of quiet period expiration. Based on these findings, the central hypothesis that this paper tests is formulated:

H2: An increase in positive (negative) sentiment score for tweets on a company increases (decreases) abnormal returns at quiet period expiration.

III. DATA AND METHODOLOGY

The key variables for analysis are a measure of abnormal returns at the end of the quiet period - derived using data from the Wharton Research Data Service Center (WRDS) - as the dependent variable and sentiment scores for a database of tweets as the explanatory variables. The analysis utilizes data from 2018 back to the JOBS act amendment in 2012. Returns are measured in real terms. The following section describes the sample data in detail and further explains the methodology employed for analysis.

III.A. Wharton Research Data Service Center

The source for abnormal quiet period returns is WRDS, an interface to a variety of datasets which target solutions for research. We use two separate datasets. The first one is Audit Analytics Initial Public Offerings, which covers U.S. registered IPOs on exchange since 2000, and is thus used to collect data on companies that IPO. The second source is CRSP, a collection of security price, return and other data, which is used to collect daily return data on each security for the duration of the quiet period.

In line with prior research on the Quiet Period, abnormal returns are measured using the variable cumulative market-adjusted returns (**CMAR**), which is measured at quiet period expiration (Bradley, Jordan & Ritter, 2003).

First the expiration date of each company was calculated by adding 40 days to the initial IPO date. The CUSIP number was also obtained from Audit Analytics, which identifies the relevant IPOs, and is used to download daily stock return data from CRSP. CUSIP is used as opposed to tickers, as they are permanent and retired stock ticker symbols are often reused. In total, after accounting for missing data, the sample consists of 225 unique IPOs and accompanying quiet period returns.

With this, the CMAR is calculated by first finding market-adjusted returns (MARs) which is the difference between the asset's return and a market index return on a given date:

$$MAR_{xd} = A_{xd} - M_{d}$$

Where MAR_{xd} is market adjusted return for firm x on day d, A_{xd} is the return of the relevant IPO and M_d is the return of the market index for firm x and day d respectively. The CRSP NYSE Value-Weighted Market Index is used as the market index for this analysis.

Consistent with prior research (Bradley, Jordan & Ritter, 2003) CMAR is then calculated using a window of 5 days surrounding the quiet period expiration date. The window consists of the expiration date and 2 days on either side of the expiration. MAR is determined for each day in the event window and CMAR is calculated as the sum of these values from day a to day b:

$$CMAR_{x}(b, a) = \sum_{t=b}^{a} MAR_{x}$$

As control variables we use data on the **AuditCompany**, which includes the auditor for each company at the time of IPO. This is following existent research which has determined that the audit quality is significantly related to post-IPO survival and returns (Badru & Zaluki, 2018). The variable **Industry** is also controlled for, using data on the NAICS (North American Industry Classification System) code of each company as industry differences are also shown to have an affect on IPO performance (Ang and Boyer, 2009)

Finally, whether a company is an Emerging Growth Company (**EGC**), defined as companies with gross annual revenue of less than $1 billion at time of IPO, is also controlled. For this, data on the gross annual revenue at the time of the IPO is collected and used to create a binary variable that accounts for whether the company is an Emerging Growth Company (EGC), defined as companies with gross annual revenue of less than $1 billion at time of IPO. This is controlled as underwriters for EGCs, as of the JOBS act, are allowed to communicate with qualified buyers both before and during the quiet period and can thus affect investor sentiment and returns. This is further warranted due to the findings of Bradley, Jordan, Ritter & Welch (2004) which showed a decrease in measured abnormal returns following regulatory changes.

III.B. A Database Of Tweets and Sentiment Scores

For the main explanatory variables we use the average **Positive Score**, **Neutral Score** and **Negative Score** for tweets on a company.

The relevant tweets are first collected and cleaned. Snscrape, a python 3 library which enables scraping social networking services, is used to obtain tweets on the IPOs during the quiet period window. The searches were limited to those in the English language. The key words for this search are the IPO company names. To prepare the data for analysis, all duplicate tweets were deleted.

The tweets were then cleaned to delete any URLS, hashtags, user handles, or emojis which are common in Twitter posts. These additions could possibly affect the final sentiment score, as analysis software only looks at words, and as thus are excluded from the sample. Using the library a sample of 38,885 tweets for the companies were finally obtained. The total number of tweets for each company was also calculated to create the variable **Number Of Tweets,** in order to conduct analysis for hypothesis 1.

For sentiment analysis, a pre-trained model that has been widely used in stock return academic literature due to its precision, recall, and ease to finetune is BERT (Hiew, Huang, Mou, Li, Wu & XU, 2019; Sousa & Sakiyama, 2019). In the main analysis TweetEVAL, a BERT-based model which has been fine tuned for Twitter Sentiment analysis is used (Baribeiri, Camacho-Collados, Neves & Espinosa-Anke, 2020).

This was chosen over a bag-of-words approach, as the latter looks at pieces of text as a collection of individual words and counts the number of positive and negative words using dictionaries (Loughran & Mcdonald, 2011) to establish sentiment. Thus, it disregards the order of the words and can lead to erroneous results if words are prefaced by other words, like conjunctions, which change their tone.

Pre-trained language models, on the other hand, have been empirically proven to be effective for improving sentiment analysis, including at the sentence level (Devlin, Chang, Lee & Toutanova, 2018). This is done by training the language models on a large corpus of text in a self supervised setting.

Using the python3 library and TweetEval, a model was coded to analyse the database of tweets while accounting for context to other words. The resulting output were the three variables: PositiveScore, NeutralScore and NegativeScore, which gives each tweet a value between 0 and 1 to reflect its positive, neutral and negative sentiment respectively. The sum of the three scores for each tweet is 1. These scores are then aggregated to indicate the Positive, Neutral and Negative Sentiment scores for each IPO in the database.

III.C. Summary Statistics

	mean	sd	min	max
CMAR	-0.0005	0.0737	-0.2399	0.3090
Number of tweets	154.5911	925.4016	2.0000	13396.0000
Positive Score	0.1859	0.1230	0.0308	0.6747
Negative Score	0.0457	0.0465	0.0075	0.4405
Neutral Score	0.7313	0.1527	0.1758	0.9325
Compound Score	0.1070	0.2139	-0.4939	0.7964
EGC	0.8933	0.3094	0.0000	1.0000
AuditCompany	10.9689	5.6580	1.0000	22.0000
Industry	8.2800	3.9626	1.0000	17.0000
Observations	225			

Table 1. Summary Statistics

Note. The sample period is 2012-2018. CMAR is calculated in a 5 day window surrounding the IPO quiet period expiration date. The Positive, Negative, Neutral and Compound Scores are results from sentiment analysis on all tweets regarding a company in the sample period. EGC is a binary variable that is 1 if total annual revenue at time of IPO is greater than 1 billion USD or 0 otherwise.

Table 1 shows the summary statistics for the sample. It can be observed that the average for the Neutral Score is 73%, which indicates that for the majority of tweets, the sentiment conveyed by the tweets are of neutral sentiment. This is followed by the Positive Score at 18% and the Negative Score at 4%. The mean of the Compound score of 0.107 indicates that on average tweets when analysed using the bag-of-words approach has a slightly positive sentiment. The value of CMAR is surprising, as it indicates that abnormal returns at the end of the quiet period is 0.07%, which is around 3 percent smaller than the values observed by Bradley et al (2003) in their seminal paper on the quiet period anomaly.

The number of tweets per company averages around 154, but what is surprising is the range of this variable, from a minimum of 2 tweets per company to a maximum of 13396 tweets per company. However, it should be noted that the maximum came from the IPO for facebook, a unicorn IPO. Finally, from the EGC variable, we can see that 89% of companies are classified as emerging growth companies. This number is consistent with research on EGCs. The mean values and corresponding summary statistics for AuditCompany and Industry do not have an interpretation as these are categorical variables.

	CMAR	Number of tweets	Positive Score	Negative Score	Neutral Score	Compound Score	EGC	AuditCompany	Industry
CMAR	1.00								
Number of tweets	-0.08	1.00							
Positive Score	-0.05	0.09	1.00						
Negative Score	-0.08	0.21	-0.21	1.00					
Neutral Score	-0.12	-0.10	-0.19	-0.08	1.00				
Compound Score	-0.09	0.17	-0.07	0.00	0.06	1.00			
EGC	-0.04	-0.23	0.06	-0.10	-0.01	-0.00	1.00		
AuditCompany	-0.00	-0.02	0.09	-0.10	0.08	0.02	0.02	1.00	
Industry	-0.06	0.07	0.20	-0.00	-0.13	-0.01	0.06	0.04	1.00

Table 2. Correlation matrix for this study

III.D. Methodology

The objective of this research is to investigate the relationship between the quiet period abnormal returns and media sentiment for companies that have recently gone through an IPO. A multivariate regression model, which measures the degree to which various independent variables and dependent variables are linearly related to each other, is used to study this. This is an extension of the classic OLS regression, which minimizes the sum of squared residuals between actual and predicted values in a model. The following section details the models and explains how effects are estimated and conclusions are drawn for analysis. For all models, the parameters $\beta_{1...6}$ measures the partial effect of the respective coefficient on CMAR. It tells us how much CMAR changes when the respective coefficient changes, ceteris paribus. εi represents the margin of error within the multivariate regression model.

To test the first hypothesis and determine the impact of the volume of tweets on abnormal returns, a simple model consisting of CMAR as the dependent variable and Number of tweets as the independent variable is used. For the Number of tweets variable, log transformations were applied as the inspected histograms concluded that they followed a lognormal distribution:

$$\text{(i) } CMAR(b,a)_i = \alpha_0 + \beta_1 * lg(\text{Number of tweets}) + \beta_2 * Industry_i + \beta_3 * AuditCompany_i + \beta_4 * EGC_i + \varepsilon_i,$$

The control variables are also included. From Audit analytics we get data on Industry, which is a categorical variable that represents the industry that the particular firm is a part of. The industries of the companies are identified using the first 2 digits of their North American Industry Classification System (NAICS) code. Similarly, the AuditCompany is a categorical variable that represents the company that conducted the audit at the time of IPO. Finally, EGC is a binary variable that is either 1 if gross annual revenue of the company is greater than 1 billion at the time of IPO and 0 otherwise.

To answer the central hypothesis - and determine the effect of media sentiment on abnormal returns, the following model with the same control variables is constructed. PositiveScore, NeutralScore and NegativeScore represent the relevant sentiment scores derived using BERT sentiment analysis.

$$\text{(ii) } CMAR(b,a)_i = \alpha_0 + \beta_1 * PositiveScore_i + \beta_2 * NeutralScore_i + \beta_3 * NegativeScore_i + \beta_4 * Industry_i + \beta_5 * AuditCompany_i + \beta_6 * EGC_i + \varepsilon_i,$$

III.E. Hypothesis Tests

The Classical Linear Regression Model assumptions are tested to see whether the model is the Best Linear Unbiased Estimator (BLUE), and to ensure that the estimated coefficients are unbiased and consistent.

For the first assumption, residuals are checked to ensure that the mean is equal to 0. If this is not true, the constant would be biased. However, this is redundant in an OLS model as regression coefficients are calculated in a way such that the residuals have a 0 mean. In line with the second assumption, homoscedasticity is tested, using a White test, to check if residuals are constant for all values of the independent variable. As the white test rejected the null hypothesis of homoskedasticity (p-value of 0.9999), robust standard errors are utilized for the regression. The third assumption states that residuals are uncorrelated. However, this assumption cannot be tested in a cross sectional dataset. The industry variable is added as a control to account for this. As it is assumed that there is no correlation between the residuals, the coefficients should be consistent.

There is no reason to believe that the fourth assumption, that independent variables are endogenous, is violated. Endogeneity can arise from attenuation bias, as a result of measurement error, which is unlikely to be present in the sample. It could arise from omitted variable bias, which is important as it suggests that the model excludes a relevant variable. A Ramsey test is performed to test for this. From the results we cannot reject the null hypothesis (p-value of 0.1804), which suggests that there is not enough evidence to conclude that the model has omitted variables. The final assumption assumes a normal distribution of residuals. This is usually not of concern due to the law of large numbers and the central limit theorem. Nonetheless, the constructed histogram for residuals does seem to follow an approximate normal distribution.

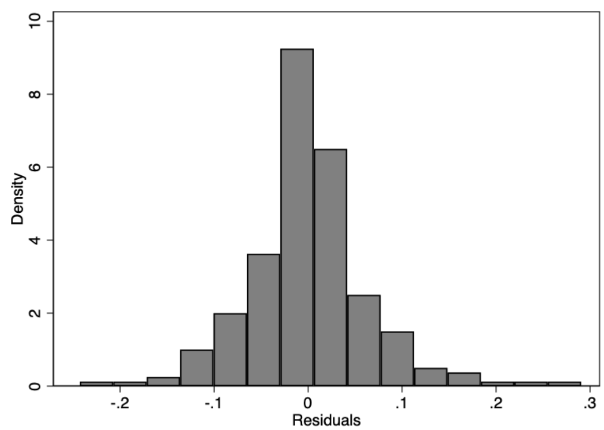

Figure 1. Plot of residuals

As all the assumptions are met, the model can be thought to have unbiased, efficient estimated coefficients. Which allows for valid hypothesis testing and implies that the model accurately studies the underlying relationship between quiet period abnormal returns and social media sentiment.

IV. RESULTS AND DISCUSSION

	(i) CMAR	(ii) CMAR	(iii) CMAR
Positive Score		-0.0588	-0.0356
		(0.0384)	(0.0633)
Negative Score		-0.1825	-0.2048*
		(0.1133)	(0.1280)
Neutral Score		-0.0721**	-0.0883**
		(0.0320)	(0.0350)
lg(Number of tweets)	-0.0017		
	(0.0053)		
EGC	-0.0004		-0.0010
	(0.0190)		(0.0192)
AuditCompany Controlled	YES	NO	YES
Industry Controlled	YES	NO	YES
Constant	0.0088	0.0715***	0.0840*
	(0.0363)	(0.0268)	(0.0472)
Observations	225	225	225
R^2	0.1525	0.0325	0.1891
Adjusted R^2	-0.0262	0.0194	0.0019

Standard errors in parentheses

Table 3. Results from regression analysis

Note. CMAR, with a window of (-2,2) is the dependent variable for all models. Model 1 includes the log Number of tweets as the independent variable, and all control variables. Model 2 includes sentiment scores from BERT as explanatory variables while Model 3 adds the controls of industry, Auditor and EGC.

$^{*}\,p < 0.10$, $^{**}\,p < 0.05$, $^{***}\,p < 0.01$

IV.A Using Volume of Tweets to explain CMAR

Table 3 shows the results of the multivariate regression models estimated on the 2012-2018 sample. Model (i) looks to answer hypothesis 1, it tests whether the quantity of tweets affects CMAR, and contrary to the expected results, and other research, we find that an increase in quantity of tweets has no effect on CMAR, with the coefficient additionally being insignificant. A 1% increase in the log Number of tweets leads to a 0.0017 unit change in CMAR. The economic magnitude of the coefficients are thus, close to 0 and insignificant.

Furthermore, the R^2 value is not sizable, with model (i) only explaining 15% of the relationship between number of tweets and CMAR. A majority of this explanatory power also comes after the inclusion of the control variable.

This overall leads to the conclusion that the Volume of tweets does not have an effect on abnormal returns at the end of the quiet period. When looking at existing literature, this is contradictory as Bushee, Cedergren & Michels (2020) concluded that greater media coverage following an IPO is associated with more purchases. This could be a reflection of the relatively smaller sample size employed in the analysis for this paper or could be due to the differences in events studied.

IV.B Using Sentiment of Tweets to explain CMAR

Column (ii) illustrates the model used to test for the second hypothesis. It indicates that the only significant variable is the neutral score, with higher neutral twitter score is significantly associated with a decrease in the value of CMAR. In Column (iii), the addition of the control variables leads to the negative sentiment score being significant, at the 10% level, with an increase in negative twitter score being associated with a decrease in the value of CMAR. The addition of the control variables of EGC, industry and audit company further strengthens the effects of the respective sentiment score. An increase in positive score by one unit, for instance, decreases CMAR by 0.0356 units in the third model and 0.0588 in the second model. This coefficient is, however, insignificant, across all models.

The economic magnitude of these coefficients are of some significance, with an increase in negative score by one unit in model (iii) leading to a decrease in CMAR value by 0.2048 units. When looking at the positive and neutral scores, however, the economic magnitude is minimal with only a decrease in 0.0356 and 0.0883 units respectively.

With regards to the control variables, contrary to expectations, the coefficient of EGC is 0.0017 when added to the model. This seems to indicate that the CMAR at the end of the quiet period is not affected if a company is classified as an EGC, however, it is hard to make a conclusion as the coefficient regarding EGC is insignificant.

When looking at the control variable of Audit Company, there are multiple coefficients which are significant, however the companies with relatively sizeable economic magnitude are audit companies: Briggs & Veselka Co, CohnReznick LLP and Moss Adams LLP, each of which leads to a decrease in CMAR. With regards to industry, there are again a few significant coefficients, but the only ones with sizable economic magnitude are manufacturing industries and scientific services.

Model (ii) only explains around 3% of the relationship, whereas model (iii) explains 19% of the relationship between twitter sentiment and CMAR. The decrease in adjusted R^2 from model (ii) to model (iii) further shows that the addition of the audit and industry controls improves the model by less than expected. These values are low when compared to similar literature in the field, such as Bajo & Raimondo (2017) whose model (despite IPO returns in general as opposed to quiet period returns) explained 33% of the studied relationship and indicate that, although significant, the sentiment scores of tweets do not explain much in the variation of abnormal returns.

With regards to conclusions for the second hypothesis, the insignificance of the Positive Score seems to indicate that positive tweets are not associated with abnormal returns. However, for Negative Score we can see a significant coefficient which proves the second part of the hypothesis by showing that an increase in negative sentiment amongst tweets does decrease the abnormal returns at quiet period expiration. This contradicts existing research again as Bajo and Raimond (2017) concluded that a more positive coverage for an IPO increases the demand in the IPO date and it generates higher abnormal returns, albeit when studying underpricing. This could be explained due to the difference in the events which are studied but understanding this is still an avenue for further research.

IV.C Additional Analysis

The robustness of the results for sentiment analysis was checked to several variations in methodology. The regression tables for these models are reported in table 4.

CMAR window of (0,2). We look at a smaller interval for the window over which CMAR is calculated. In the original analysis CMAR was calculated over a window of 2 days prior to the quiet period expiration and 2 days after. The window of time solely after the expiration has been studied before (Bradley et al 2004), and thus a robustness check was warranted. The results led to neutral scores being insignificant but negative scores still being significant. Thus, with regards to the hypotheses, the results are still robust.

Including number of followers. Although it has not been included in prior literature, the reach of a particular tweet is determined by the number of followers the user has. Thus, using Snscrape, we derive this number and add it to our regression. The results do not change much after this addition.

Table 4 summarizes the results from the regression analysis conducted for the first two robustness checks.

	(1)	(2)
	CMAR(0,2)	CMAR
Positive Score	-0.0019	-0.0387
	(0.0500)	(0.0645)
Negative Score	-0.1560*	-0.2015
	(0.0896)	(0.1289)
Neutral Score	-0.0572	-0.0933**
	(0.0379)	(0.0358)
lg(Number of tweets)	-0.0054	-0.0005
	(0.0051)	(0.0057)
EGC	-0.0094	0.0039
	(0.0112)	(0.0200)
AuditCompany Controlled	YES	NO
Industry Controlled	YES	NO
Followers		-0.0000
		(0.0000)
Constant	0.1508***	0.0785
	(0.0473)	(0.0528)
Observations	225	215
R^2	0.1977	0.1872
Adjusted R^2	0.0126	-0.0054

Table 4. Regression analysis for robustness checks

Note. CMAR, with a window of (0,2) is the dependent variable for Model 1. Model 2 shows the results for analysis with the regular CMAR window but with the addition of the Number of followers as a control.

Standard errors in parentheses * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

VADER, A Bag-Of-Words Approach: Despite research empirically testing and concluding that Roberta software for sentiment analysis is better than Bag-of-words (Devlin, Chang, Lee & Toutanova, 2018), there exists a line of research which has used the bag-of-words method in their analysis of returns. Hajek (2016), for instance, used bag-of-words analysis to find that there is a strong relationship between sentiment in annual reports and abnormal stock return data. One tool used in academia for bag-of-words analysis is VADER, with Pano & Kashef (2020) using it to anlayse tweets on Bitcoin during Covid-19, finding that there is a significant short-term correlation between prices and daily tweet sentiment.

Due to this, this section looks into a methodological variation and conducts the same analysis, for the same sample of tweets with CMAR as the dependent variable and the same control variables, using VADER - a bag of word approach - to establish sentiment score. As with Pano & Kashef (2020), we used the compound sentiment - a score that weights the positive, negative, and neutral sentiment in each tweet - as the independent variable. The individual positive, negative and neutral scores that VADER calculates were also briefly looked into, but yielded seeming anomalous and biased results (see Appendix B).

$$\text{(iii) } CMAR_x(b,a)_i = \alpha_0 + \beta_1 \text{*Compound Score} + \beta_2 \text{*Number of tweets} + \beta_3 \text{*Industry}_i + \beta_4 \text{*AuditCompany}_i + \beta_5 \text{*EGC}_i + \varepsilon_i,$$

The results from the regression are shown in table 5.

	(1)
	CMAR
Compound Score	-0.0485**
	(0.0204)
EGC	0.0014
	(0.0191)
AuditCompany controlled	YES
Industry controlled	YES
Constant	0.0118
	(0.0239)
Observations	225
R^2	0.1683
Adjusted R^2	-0.0070

Table 5. Results from VADER regression analysis

Note. CMAR, with a window of (-2,2) is the dependent variable for the model. Model 1 includes the sentiment score from VADER as explanatory variables and the controls of industry, Auditor and EGC.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

From the results, the coefficient of -0.05 for the Compound Score variable - which is significant at the 5% level - is of interest. The compound score is a value between -1 and +1, with -1 indicating that a tweet has a completely negative sentiment, and +1 indicating that a tweet has a completely positive sentiment. The coefficient of -0.0485 thus indicates that an increase of the compound score by 1 unit decreases CMAR by 0.0485 units and conversely that a decrease of compound score by 1 unit increases CMAR by 0.0485 units. This, thus, also indicates that an increase in positive sentiment decreases CMAR and an increase in negative sentiment increases CMAR, ceteris paribus. This is conflicting to the existing research and formulated hypothesis and can be explained by the framework which the bag-of-words approach employs for its analysis.

The difference in R^2 between the VADER and BERT models is around 2 percent. To better illustrate which model fits the data better, the Akaike Information Criterion (AIC) and the Bayesian Information criterion (BIC) for both the BERT and VADER models are calculated. The results are shown in table 5.

	(1) BERT	(2) VADER
AIC	-524.8187	-523.2142
BIC	-425.7518	-430.9795

Table 5. AIC and BIC scores for the two models

As shown in current research (Ludden, Beal & Sheiner, 1994) the selection criteria for the AIC enables the ability to select the correct model. The lower the AIC and BIC scores, the better. In our results, we can see that although the VADER model has a lower AIC score, albeit only by around 1 point, the BERT model is shown to have a lower BIC score, by a larger margin. The results of this test combined with the R^2 , puts further emphasis on the relative strength of an NLP approach to the bag-of-words approach.

V. CONCLUSION

The study investigates the relationship between social media sentiment during the quiet period in an attempt to explain the abnormal returns at the end of the quiet period. This paper sheds some light on the phenomena as although IPOs have been a topic with a plethora of research, discussion regarding the quiet period itself and factors that contribute to the anomaly have not been studied extensively.

The paper contributes to finance literature on the IPO quiet period by analyzing all tweets on a sample of IPOs from 2012 to 2018, using both Natural language processing and a bag of words approach, and then conducting a multivariate regression with audit company and industry as control variables. The results indicate that the volume of tweets does not have an impact on abnormal returns. It further indicates that a higher neutral and negative sentiment score are significantly associated with a decrease in abnormal returns. Positive sentiment scores, however, do not have significant impact. The addition of the respective control variables strengthens the effect, but whether or not a company is an EGC has no impact. When looking at a bag-of-words approach, we find conflicting results: an increase in positive sentiment decreases CMAR and an increase in negative sentiment increases CMAR, ceteris paribus.

From the results, individuals who tend to look into media as a source of information for retail investing can make better and more informed decisions when interpreting Tweets and the tone conveyed in them. Although both models presented different results, it can be concluded that investors should pay special attention to companies for which aggregate sentiment on twitter is negative or neutral

Prior discussion on the quiet period establishes a concern over information outside a company's prospectus influencing retail trading decisions (Bushee et al, 2020). Concern was also shared regarding the control of media, which is out of the scope of the SEC. Although this continues to be the case, the low economic magnitude of the coefficients for sentiment, suggest that no regulation changes with reference to the media during the quiet period is needed. This is further supported by the fact that CMAR is much lower now compared to early research in the field, which suggests that the existing changes in regulations have already played its part.

Although the analysis has contributions to the existing literature, there are multiple areas for future research. The low R^2 and explanatory power of the models suggest that there are other factors which can have an impact on abnormal returns. Some may include, for instance, the number of retweets the tweets have. The plausibility of this, however, is called into question by the recent account restrictions by Twitter to discourage data scraping (Saligrama, 202). This also brings to question the relevance of twitter as an information source going forward. To further explore the effect that media has on the Quiet Period Anomaly, similar analysis with newspaper sentiment as an explanatory variable is also warranted due to the significant amount of market participants who use this as a source of information. Overall, this paper is only the first step to comprehending the anomaly that is the Quiet Period.

BIBLIOGRAPHY

Bajo, E., & Raimondo, C. (2017). Media sentiment and IPO underpricing. *Journal of Corporate Finance, 46,* 139–153.

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

Barbieri, F., Camacho-Collados, J., Neves, L., & Espinosa-Anke, L. (2020). Tweeteval: Unified benchmark and comparative evaluation for tweet classification. *arXiv preprint arXiv:2010.12421*.

Badru, B. O., & Ahmad-Zaluki, N. A. (2018). Explaining IPO initial returns in Malaysia: ex ante uncertainty vs signalling. *Asian Review of Accounting*.

Bradley, D. J., Jordan, B. D., & Ritter, J. R. (2003). The quiet period goes out with a bang. *The Journal of Finance*, *58*(1), 1-36

Bradley, D. J., Jordan, B. D., Ritter, J. R., & Wolf, J. G. (2004). THE IPO QUIET PERIOD REVISITED.

Bushee, B., Cedergren, M., & Michels, J. (2020). Does the media help or hurt retail investors during the IPO quiet period? *Journal of Accounting and Economics, 69*(1), 101261.

Carr, C. T., & Hayes, R. A. (2015). Social media: Defining, developing, and divining. *Atlantic journal of communication*, *23*(1), 46-65.

Carter, R. B., Dark, F. H., & Singh, A. K. (1998). Underwriter Reputation, Initial Returns, and the Long-Run Performance of IPO Stocks. *The Journal of Finance, 53*(1), 285–311.

Cedergren, M. C. (2014). Joining the Conversation: How Quiet Is the IPO Quiet Period? (SSRN Scholarly Paper No. 2536004).

Chowdhury, G. G. (n.d.). Natural Language Processing.

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

DellaVigna, S., & Kaplan, E. (2007). The Fox News effect: Media bias and voting. *The Quarterly Journal of Economics*, *122*(3), 1187-1234.

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

Fung, G. P. C., Yu, J. X., & Lam, W. (2002). News sensitive stock trend prediction. In *Advances in Knowledge Discovery and Data Mining: 6th Pacific-Asia Conference, PAKDD 2002 Taipei, Taiwan, May 6–8, 2002 Proceedings 6* (pp. 481-493). Springer Berlin Heidelberg.

Gentzkow, M., & Shapiro, J. M. (2006). Media bias and reputation. *Journal of political Economy*, *114*(2), 280-316.

Greitemeyer, T. (2011). Effects of prosocial media on social behavior: When and why does media exposure affect helping and aggression?. Current Directions in Psychological Science, 20(4), 251-255.

Hiew, J. Z. G., Huang, X., Mou, H., Li, D., Wu, Q., & Xu, Y. (2019). BERT-based financial sentiment index and LSTM-based stock return predictability. *arXiv, 1906.09024*.

Keim, D. B. (1983). Size-related anomalies and stock return seasonality: Further empirical evidence. *Journal of financial economics*, *12*(1), 13-32.

Kwan, T. (2015). Twitter Volume and First Day IPO Performance. New York: The Leonard Glucksman Institute for Research in Securities Markets, Stern School of Business

Liew, J. K.-S., & Wang, G. Z. (2016). Twitter Sentiment and IPO Performance: A Cross-Sectional Examination. *Journal of Portfolio Management, 42*(4), 129–135.

Loughran, T., & McDonald, B. (2011). When is a liability not a liability? Textual analysis, dictionaries, and 10‐Ks. *The Journal of finance*, *66*(1), 35-65.

Loughran, T., & Ritter, J. R. (1995). The new issues puzzle. *The Journal of finance*, *50*(1), 23-51.

Medhat, W., Hassan, A., & Korashy, H. (2014). Sentiment analysis algorithms and applications: A survey. *Ain Shams engineering journal*, *5*(4), 1093-1113.

Neuhierl, A., Scherbina, A., & Schiusene, B. (2013). Market Reaction to Corporate Press Releases. The *Journal of Financial and Quantitative Analysis, 48*(4), 1207–1240.

Pano, T., & Kashef, R. (2020). A complete VADER-based sentiment analysis of bitcoin (BTC) tweets during the era of COVID-19. *Big Data and Cognitive Computing*, *4*(4), 33.

Papers with Code—Financial_phrasebank Benchmark (Text Classification). (n.d.). Retrieved April 2, 2023, from https://paperswithcode.com/sota/text-classification-on-financial-phrasebank

Ritter, J. R. (1991). The long‐run performance of initial public offerings. *The journal of finance*, *46*(1), 3-27.

Ritter, J. R., & Welch, I. (2002). A review of IPO activity, pricing, and allocations. *The journal of Finance*, *57*(4), 1795-1828.

Saligrama, A. (2023, July 1). Musk says Twitter will limit how many tweets users can read. Reuters. https://www.reuters.com/technology/musk-says-twitter-applies-temporary-limit-address-data-scraping-system-2023-07-01/

Shen, D., Urquhart, A., & Wang, P. (2019). Does twitter predict Bitcoin?. Economics Letters, 174, 118-122.

Sousa, M. G., Sakiyama, K., de Souza Rodrigues, L., Moraes, P. H., Fernandes, E. R., & Matsubara, E. T. (2019, November). BERT for stock market sentiment analysis. In *2019 IEEE 31st International Conference on Tools with Artificial Intelligence (ICTAI)* (pp. 1597-1601).

The JOBS Act After One Year: A Review of the New IPO Playbook. (n.d.).

Thomson, D. (n.d.). *Update on emerging growth companies and the Jobs Act*. PwC.

https://www.pwc.com/us/en/services/consulting/deals/emerging-growth-companies-jobs-act.html
#:~:text=More%20than%2080%25%20of%20all,available%20under%20the%20JOBS%20Act.

APPENDIX

Appendix A: VADER analysis extended

This section briefly looks into the results of using VADER when scores are separated to positive, negative and neutral. As VADER utilizes a bag of words approach, these scores are calculated by looking into the number of positive, negative and neutral words in a tweets. As such, the results are biased as the analysis does not consider prepositions or the overall meaning of the tweet.

This can be seen through the magnitude of the coefficients for the scores, all in the 20s. This seems to indicate a strong association with CMAR as it suggests that an increase in 1 unit of the scores will increase CMAR by around 20 percentile points, ceteris paribus. However, none of these scores are significant when using this method of analysis, and as such we cannot derive any concrete conclusions.

	CMAR
Vader Positive	24.82
	(29.16)
Vader Negative	25.27
	(29.25)
Vader Neutral	24.94
	(29.16)
EGC	0.00
	(0.02)
AuditCompany Controlled	YES
Industry Controlled	YES
Constant	-24.92
	(29.16)
Observations	226
R^2	0.17

Table 7. Results from Vader regression analysis

Note. CMAR, with a window of (-2,2) is the dependent variable for the model. Model i includes the Positive, Negative and Neutral scores from a Bag-of-words approach, and all control variables. Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.