ERASMUS UNIVERSITY ROTTERDAM
ERASMUS SCHOOL OF ECONOMICS
MSc Economics & Business or
Specialization Financial Economics

Investor Attention and Healthcare Stocks before and during Covid-19

Author: S. D. van Hövell van Wezeveld en Westerflier
Student number: 623620
Thesis Supervisor: Dr. J. J. G. Lemmen
Second Reader: Dr. J. C. M. Kil
Finish date: 12-07-2022

Preface and Acknowledgement

I would like to express my deep gratitude to my master thesis supervisor, Dr. J. J. G. Lemmen. His helpful insights, supportive and useful critiques, and expertise within finance have provided me with great support throughout the entire process.

Abstract

This thesis studies the effect of investor attention, measured by standardized Google search volumes, on stock returns of healthcare companies from the S&P 500 and S&P 400 mid-cap during Covid-19. The exact focus of this study is on the difference in effect between the period before and period after the Covid-19 outbreak. This is the first study that combines investor attention, using Google searches for specific company names and company tickers, with stock returns of the US healthcare sector and the Covid-19 pandemic. Therefore, this study contributes to the existing literature on the effect of investor attention, measured by GSV data, on stock performance during periods of crisis. This research is based on a panel data set covering the period from January 2019 to April 2021. Moreover, three measures of Google search volumes (by ticker, by company name, and combined), stock data from the CRSP database, and a set of control variables from the Compustat database (both WRDS) were used to conduct this research. In addition, the results were obtained using OLS and Fixed effects regressions, including firm/subsector and week fixed effects. First of all, a reversal point was established, based on the fact that the effect of SGSV data on stock returns of US healthcare companies is positive in the short term (1 to 2 weeks) and negative in the longer term (after reversal point). The main finding of this thesis is that the outbreak of Covid-19 has an amplifying effect on the positive short-term (1 to 2 weeks) and negative long-term (after reversal point) effect of investor attention on stock returns of healthcare companies from the S&P 500 and 400. In addition, companies within the sample that are directly related to the Covid-19 pandemic received more attention than the other companies. However, the results regarding the effect of investor attention for directly/not directly related companies are ambiguous. Lastly, firm size appeared to have a positive effect on the effect of investor attention on stock returns during the Covid-19 period.

Table of contents

List of Tables

1. Introduction

The date was March 11, 2020, when Covid-19 was officially recognized as a global pandemic. On this day it was officially confirmed by the World Health Organization (Zhang et al., 2020). Not much later, more than 170 countries were affected by Covid-19 which led to the fact that about 1/3 of the world's population had to deal with strict policies (Xu et al., 2020). These policies had major economic consequences, because a large part of economic traffic was brought to a standstill. Financial markets were hit and this was reflected in stock returns and stock volatility all over the world. On March 23, 2020, the S&P 500 index had fallen 34% from its peak a month earlier on February 19 (Demers et al., 2020).

A specific concept that can influence the return and volatility of stocks is investor attention, as studied by Barber and Odean (2008). Individual investors tend to buy stocks that grab attention. They showed that stocks that make headlines, stocks with a high abnormal trading volume, and stocks with extraordinary 1-day returns, are bought more frequently by individual investors. This phenomenon is central to this study. In addition, Cao et al. (2009) show that the attention of (individual) investors is limited to a number of assets with which they are already familiar. It is impossible for individual investors to consider every investment option in their decision. The Covid-19 crisis offers a unique opportunity to examine the impact of investor attention on stocks during a recent global health crisis.

Given the health nature of the Covid-19 crisis, this offers an opportunity to investigate the effect on healthcare companies in particular. Therefore, the concept of investor attention described above could apply to healthcare companies. During this crisis, these companies were in the news more often and this also reached the individual investor. Therefore, the focus of this thesis is specifically on healthcare companies. In addition, a distinction is made between companies directly related to the virus and healthcare companies not directly related to Covid-19. It is expected that not all healthcare companies will automatically be in the news more often because of this crisis. However, this is expected to apply to companies that are related to the development or production of vaccines, for example.

The impact of investor attention has been widely studied in the existing literature and is still a much-researched concept. Several studies examined whether data from Google Trends can be used as a forecast for stock returns (Bank et al., 2011; DA et al., 2011; Preis et al., 2013; Bijl et al., 2016). In other words, they investigated whether investor attention has an influence on stock prices and whether these stock prices will increase or decrease. Another research by Blitz et al. (2019) studies whether the investor attention hypothesis, based on media attention, can potentially explain the volatility effect. The existing literature agrees on the fact that high Google Search Volumes (GSV) will lead to an increase of the stock returns on the short term (first 1-2 weeks), followed by a reversal point after a short time period (Bank et al., 2011; DA et al., 2011; Preis et al., 2013; Bijl et al., 2016). Although, the results on the magnitude and timing of the reversal are contradictory. These studies focus on individual investors, who are sensitive to high profile stocks as described by Barber and Odean (2008).

Institutional investors continuously monitor stocks and are less influenced by these eye-catching stocks. This may also explain the reversal that occurs after a temporary rise in stocks following high Google Search Volumes, as Institutional investors sell. Institutional investors rely more on professional data rather than Google Trends, such as Bloomberg (Ben-Rephael et al., 2017). Therefore, the focus and scope of this thesis is on the individual investor. In addition, high Google search volumes are positively related to high stock price volatility (Bank et al., 2011; Vlastakis & Markellos, 2012; Dimpfl & Jank, 2015; Kim et al., 2019). However, the effect of investor attention during periods of crisis, such as the Covid-19 period, has been less studied. Wan et al. (2021) examined the effect of investor attention on the performance of clean energy firms versus fossil fuel firms during Covid-19. Covid-19 has had a major impact on the entire financial world, including the stock market, providing an opportunity to further study the impact of investor attention. Smales (2020) examined the effect of investor attention on US stock prices during Covid-19 per sector, with a total of 11 sectors. Although the stock returns of most sectors have been negatively affected by investor attention during Covid-19, several sectors have benefited from the crisis. One of the benefiting sectors is the healthcare sector, which is examined in the current study. In contrast, the existing literature focus on the Google Trends search terms related to 'Covid-19' or 'Coronavirus' (Smales, 2020; Smales, 2021; Iyke & Ho, 2021; Vasileiou, 2021; Dey et al., 2022), while the current research focuses on individual companies through company names and/or ticker. Therefore, the research question of this study is the following:

'What is the effect of GSV on healthcare stocks returns during Covid-19 compared to the pre-Covid period?'

The results of this research add to the existing literature in several ways. First of all, it contributes to the literature on predicting stock market returns by Google Trends. Secondly, this research adds to the results on the effect of investor attention during crisis periods, especially Covid-19. During Covid-19, healthcare firms and their stocks obviously receive more attention due to the health nature of the crisis. Investor attention is a widely researched concept that shows different results in specific situations. The most recent (health) crisis, the Covid-19 crisis, offers a unique opportunity to study the effect of investor attention in a crisis situation. In addition, Covid-19 offers an opportunity to study the healthcare sector during a crisis period, given the health nature of the crisis. Thirdly, most literature on investor attention during Covid-19 focusses on Google Search Volumes of 'Coronavirus' or 'Covid-19'. These studies examine the relationship between the fear of Covid-19 and the effect on certain stocks or indices. The current study will focus on Google search terms of the individual company tickers and company names. Fourthly, the current study will focus on the US stock market and will draw a comparison between the period before and during the Covid-19 crisis.

The main findings of this thesis consist of an established reversal point after 1 to 2 weeks and an amplifying effect of Covid-19 on the effect of investor attention on stock returns of healthcare companies from the S&P 500 and 400. This means a more positive effect in the short term (first 1 to 2 weeks) and a more negative effect in the longer term (after reversal point), compared to the period before Covid-19. Finally, firm size and relatedness to the virus were also found to be positively correlated with the amount of investor attention the healthcare companies received.

The remainder of this research will be structured as follows: The existing literature on the healthcare sector, investor attention, Covid-19 and investor attention in periods of crisis (financial crisis and Covid-19 crisis) will be discussed in chapter 2. Based on this, the hypotheses will be formulated. Subsequently, chapter 3 describes which data is used and where it is extracted from. Thereafter, the methodology will be discussed in chapter 4 and the results can be found in chapter 5. Finally, Chapter 6 includes the discussion and chapter 7 will provide an overall conclusion to the results and this thesis.

2. Literature review

The fact that people unconsciously focus on certain information was already established in 1973. Kahneman (1973) described, among other things, that a person can consciously process only a small part (limited) of all information. This can be applied to several theories on financial markets. Therefore, it can be concluded that this is in contrast to Fama (1965) who described the well-known efficient market hypothesis. This hypothesis concludes that all available information is included in the stock price. New information will be immediately incorporated into the stock price. However, the existing literature shows several anomalies from this theory. Merton (1987) elaborates on the fact that people only have limited attention. He claims that the prices of stocks are imperfectly diversified for this reason. His conclusion is that companies that receive little attention should deliver higher returns to compensate for the aforementioned imperfect diversification. That is why companies that receive a lot of attention will experience a price increase in the short term, which will be followed by lower returns in the longer term. Merton thus already demonstrated the phenomenon of price pressure. Hong and Stein (1999) Found that new information is not immediately fully processed by the markets. This process is gradual which means that not all information is processed in the markets and stock prices at all times. Later, Hong and Stein (2007) also appeared to disagree with the efficient market hypothesis. They concluded that financial markets cannot always behave completely rationally. This research focuses on one of the deviations from this theory.

Investor attention is a concept that is scientifically researched and proven for the first time by Barber and Odean (2008). This is called the investor attention hypothesis and implies that people tend to buy stocks that stand out in the news, stock that have extraordinary 1-day returns, or stocks that have a high abnormal trading volume. Individual or retail investors do not have the resources to consider every possible stock or other investment in their investment choice. Therefore, these investors mainly buy stocks that have caught their attention. As a result of the research by Barber and Odean (2008), the effect of investor attention on several concepts has been extensively studied. The existing literature on investor attention mostly focuses on the effect on share prices and the volatility of these prices. The recent Covid-19 crisis period provides an opportunity to investigate the effect of this phenomenon on stock performance in times of crisis. Hence, a comparison with the pre-Covid period will be examined.

2.1 Literature on healthcare sector in the US

The first section of the literature review focuses on the healthcare market in the US, as this is a diverse and broad market. First of all, there is a distinction between the private and public healthcare sector. This thesis only focuses on the public healthcare sector, as there is a lack of data availability of the private healthcare companies. In general, the private healthcare firms are smaller for-profit firms that are not funded by the government. The public healthcare sector contains larger and more well-known companies. However, the private and public sector are closely linked together. Sheinson et al. (2020)

examined the impact of both the private and the public healthcare sector on the US healthcare resource utilization during Covid-19. Both sectors cooperated to fight de virus. In April 2020, for example, they have set up a collaboration (the Accelerating COVID-19 Therapeutic Interventions and Vaccines) with both the public and private sectors. The ACTIV partnership has been created, among other things, to contribute and facilitate the development of vaccines. This partnership is a good example of the cooperation between the private and public healthcare sector during Covid-19. Outside of Covid-19, these kinds of partnerships, called PPPs (public-private partnerships), are common in the US healthcare sector (Strasser et al., 2021). Especially during times of health crisis, such as the Covid-19 pandemic, these collaborations are valuable (like the ACTIV partnership). The Covid-19 outbreak caused a huge demand for several products during Covid-19 (Vrontis et al., 2021). He mentions, among other things, the urgent need for hospitals, their equipment, and care. There is an enormous demand for products such as face masks, other personal protection, beds for intensive care, means for disinfection and ventilators. This puts enormous pressure on the healthcare sector, which made collaborations between the public and private sectors even more important.

Next to the public and private healthcare sector, there are many subsectors within the healthcare sector. Abraham et al. (2015) discuss the fact that the US healthcare sector consists of several subsectors, covering a wide range of different services. Examples include drug companies (pharmaceuticals and biotechnology), hospitals, and suppliers of medical equipment. One of the better performing subsectors is the pharmaceutical industry, characterized by a series of acquisitions. A common phenomenon is the fact that the large established companies from the public healthcare sector take over the promising and innovative companies from the private sector. The US healthcare sector is known for being a suitable sector for acquisitions and private equity. Abraham et al. (2015) describe how investments in the healthcare sector are more profitable for both buyers and sellers in the US, compared to other sectors. However, the valuation of these acquisitions and healthcare companies depends on both tangible and intangible assets. Compared to other sectors, a large part of the total assets is covered by these intangible assets (Rider et al., 2018). This is typical for the healthcare sector. Moreover, the complexity to make accurate estimates of the value of (private) healthcare companies therefore increase. Based on a report by Reilly (2013), intangible assets exist in many different forms within the healthcare sector. For example: patient relationships, goodwill (both personal and for the entire entity), supplier purchase agreements, and licenses. Subsequently, Rider et al. (2018) identify intangible assets such as human capital, research, and partnerships that are generally more difficult to value. Therefore, an accurate estimate of intangible assets is often complicated.

In addition to the study by Abraham, several studies have focused on the (positive) impact of Covid-19 on healthcare stocks. Oncu (2021) examines the biotechnology sector, a subsector of the healthcare sector, and finds positive effects on stock performance. He focuses on companies that developed drugs or vaccines against the virus. According to them, every crisis provides opportunities for certain groups or sectors. He et al. (2020) also found that the healthcare sector in general benefits

from the health-related crisis. In addition, the results of another study by Mazur et al. (2020) show that the healthcare sector is one of the sectors with positive returns from the March 2020 crash.

Table 1. Overview of literature on the healthcare sector

Author(s) (Publication year)	Time period	Region	Method	Results
Reilly (2013)	2012-2013	US	Descriptive research	Healthcare intangible asset valuation matters
Abraham et al. (2015)	1995-2004	US	Performance analysis	Healthcare sector is good to private equity and specialist firms have an edge over generalists.
Rider et al. (2018)	2017-2018	US	Experimental	Knowledge and skills are important valuable factors
Sheinson et al. (2020	2020-2021	US	Diagnostics testing and treatments	Both sectors worked together to reduce resource utilization
He et al. (2020)	2020	China	Event study	Healthcare sector one of the sectors that benefit from the crisis
Mazur et al. (2020)	2020	US	Event study	One of the sectors with positive returns during market crash was healthcare sector
Strasser et al. (2021)	2018	US	Semi-structured interviews	PPPs are extremely important in times of crisis like Covid-19
Oncu (2021)	2020	Worldwide	Fama and French, SAAR	Positive effect on healthcare sector Biotechnology
Vrontis et al. (2021)	2020-2021	US	Survey for hospitals	dynamic managerial innovative practices have a positive effect on competitive advantages and non-financial performance

2.2 Literature on investor attention (measured by GSV/Baidu)

Several studies examined the effect of investor attention by analyzing Google Search Volumes (GSV). Previous literature on investor attention all use some form of Google Trends (GT), Google Search Behavior (GSB), or Search Volume Indexes (SVI) as a proxy to examine the effect of investor attention on the stock market (Bank et al., 2011; Da et al., 2011; Preis et al., 2013; Ying et al., 2015; Bijl et al., 2016). According to these researches, stock prices appear to be predictable to a certain extent in both the short and long term by using Google Search Volumes. Bank et al. (2011) find an increase in search volumes shows a positive relationship with both trading activity and stock liquidity, but also has a temporarily positive effect on the stock price in the short term. Da et al. (2011) find similar results and conclude that a price reversal takes place after 1 to 2 weeks. In contrast, Ying et al. (2015) conducted a similar study but focused on the Chinese market, using Baidu search data instead

of Google Search. Although the results are broadly similar, they found a shorter period before the reversal point occurred and concluded that the previously observed increase could not be completely offset within one year. However, Bijl et al. (2016) focus on individual firms from the S&P 500. They find, in contrast to the previous mentioned studies, a much faster or immediate price drop after high Google Search Volumes. Therefore, they notice that scholars such as Da et al. (2011) base their results on an earlier period (2004-2008), which can influence the results. From this, it can be deduced that the impact of this relationship can differ over time, partly due to the development of financial markets and technology. More recent literature focuses on continents outside US and Europe. Yoshinaga and Rocco (2020) focus on the Brazilian stock market from 2014 to 2018 and have found a negative relationship. Nguyen et al. (2019) study 5 emerging markets, namely Malaysia, Thailand, Philippines, Indonesia, and Vietnam. They find conflicting results for different countries, as Google Searches, depending on the emerging country, have a positive or negative effect on the stock price. Therefore, it can be concluded that the existing literature does not show consensus on the effect of investor attention (GSV) on stock performance. Although, most studies conclude a short period of an increase in stock price, followed by a reversal point and decrease.

Table 2. Overview of literature on GSV data

Author(s) (Publication year)	Time period	Region	Method	Results
Barber & Odean (2008)	1991-1996	US	Time Series	Individual investors buy attention grabbing stocks
Cao et al. (2009)	1975-2006	Germany, Japan, UK, US	Calibration analysis	Familiarity, under diversification, home- and local biases
Bank et al. (2011)	2004-2010	Germany	Panel regression + time series	Temporarily higher returns + increased stock volatility and trading activity
Da et al. (2011)	2004-2008	US	Google searches and Russel 3000	New measure of investor attention
Vlastakis & Markellos (2012)	2004-2009	US	Google Trends, GARCH, Granger Causality	variations in info demand have an effect in terms of volatility and trading volume
Preis et al. (2013)	2004-2011	US	Google Trends	Increases in Google search volumes for keywords relating to financial markets before stock market falls

Dimpfl & Jank (2015).	2006-2011	US	VAR model	Search volumes Granger-cause volatility
Ying et al. (2015)	2006-2011	China	Baidu index	Investor attention has a positive effect on stock returns in the first week, after 2-4 week a reversal
Bijl et al. (2016)	2008-2013	US	Standardized Google Searches	High GSV leads to negative returns
Nguyen et al. (2019)	2009-2016	5 Asian countries	Google searches, Fama-French model	Effect of investor attention differs per country
Blitz et al. (2019)	2001-2018	US	Sorting approach	Volatility effect not only declared by attention in media
Swamy et al. (2019)	2012-2017	India	GSVI	Higher GSVI predicts higher returns in first 1 or 2 weeks
Kim et al. (2019)	2012-2017	Norway	Google Searches	Google Searches is related to stock volatility but not to stock returns

2.3 Literature on investor attention during other crises

The last crises before the Covid-19 crisis were the financial crisis of 2007-2008 and the subsequent Eurozone crisis of 2009-2010. During these periods, the effect of investor attention has also been studied. Yu & Hsieh (2010) examined the financial crisis and the Taiwan stock exchange. In this study they looked at different investors (individual and institutional) and their investment behavior based on attention grabbing factors. They concluded that the financial crisis affected purchasing behavior based on investor attention. During the crisis of 2007, investors appeared to act less emotionally and are therefore less sensitive to striking stocks. In addition, Peltomäki & Vähämaa (2015) investigated investor attention during the Eurozone crisis that followed the financial crisis. In this research she focused on the attention that the Eurozone crisis received, based on GSV, and how this influenced the herding behavior within the banking sector in various European countries. Like the existing literature on investor attention during Covid-19, which will be discussed below, the research focuses on search terms related to the crisis itself. This is in contrast to the current study in which the focus is on search terms related to company ticker and company name.

2.4 Literature on investor attention during Covid-19

Despite the short existence of Covid-19, a few papers in the existing literature already investigated the effect of the crisis on investor attention and stock performance. Several studies examined the effect of investor attention during Covid 19 on stock volatility (Jiang et al., 2021; Wang et al., 2021) and stock

prices (Smales, 2020; Iyke & Ho, 2021; Vasileiou, 2021; Dey et al., 2022; Padungsaksawasdi & Treepongkaruna, 2021). They all conclude that investor attention for the coronavirus has a positive relationship with stock volatility and a relationship with stock prices in both the short and long run. In addition, these studies all use Search Volumes or Baidu to measure the degree of investor attention. However, the searches include terms such as 'coronavirus' and 'Covid-19'. Therefore, their focus is on the relationship between attention to the virus and its effect on stock performance. The existing literature thus leaves a gap to measure the influence of investor attention for specific companies during the Covid-19 period. Smales (2020) examined the effect on 11 different industries and found that Covid-19 attention has had a positive effect on several industries, including the healthcare sector.

Table 3. Overview of literature on Investor Attention during crises periods

Author(s) (Publication year)	Time period	Region	Method	Results
Demers et al. (2020)	2020	US	Multiple regression analysis	ESG did not immunize stocks during the COVID-19 crisis, but that investments in intangible assets did.
Smales (2020)	2019-2020	US	GSV based on Da et al (2015), Studied 11 different sectors	heightened attention towards COVID-19 negatively influences US stock returns
Xu et al. (2020)	2019-2020	China	Baidu index as during covid-19	Chinese stock returns lower after outbreak
Yoshinaga & Rocco (2020)	2014-2018	Brazil	GSV	Increase in GSV is followed by lower returns
Zhang et al. (2020)	2020	Several countries	Correlation analysis	Volatility increases due to Covid-19 outbreak
Smales (2021)	2020	G7 and G20 stock markets	GSV as proxy investor attention (fear)	Negatively influenced stock returns during Covid-19
Wan et al. (2021)	2019-2020	China	Fixed effects regression	COVID-19 outbreak negatively impacted both clean energy and fossil fuel firm
Jiang et al. (2021)	2020-2021	China	Baidu index as measure for investor attention, GARCH	Forecast model including the Baidu index is significantly better than the benchmark model

Iyke & Ho (2021)	2020	Africa	EGARCH	Effect of investor attention differs per country
Padungsaksawasdi & Treepongkaruna (2021)	2020	71 countries	GSV, OLS	SVI and the growth in confirmed cases lower the global stock market returns
Vasileiou (2021)	2019-2020	10 developed countries	GSVI, EGARCH model	GSVI enables to draw info regarding the impact of the COVID-19 fear on stock market performance
Dey et al. (2022)	2020	US	Google searches for fear Covid-19	COVID-19 cases and deaths, its local spread, and Google searches have impacts on abnormal stock prices

2.5 Literature on Granger Causality tests

Several of the studies mentioned in the meta tables above performed Granger Causality tests, which will also be performed in this thesis. In this section, the results of the Granger Causality tests in these studies will be discussed. These tests have to be performed to draw a conclusion about the causality between the dependent and independent variable. In this thesis, the Granger Causality tests examine the causality between investor attention and stock returns. It is tested whether investor attention, measured by GSV, Granger-cause stock returns or whether stock returns Granger-cause investor attention. This should be tested to determine how to specify the regression. Thereafter, the results can be used to determine whether there is a simultaneity problem, a form of endogeneity problems. Roberts & Whited (2011) named four types of endogeneity problems: omitted variables, simultaneity, selection bias, and measurement errors. The simultaneity problem can occur when the dependent (stock returns) and independent (GSV) variables influence each other. In this study, the expected result is that investor attention Granger-cause stock returns.

Vlastakis & Markellos (2012) examined the relationship between information demand and information supply. Information demand is measured by Google Searches Volumes from Google Trends. The information supply is reflected by data from the Reuters NewsScope Archive, which is a large news provider. To conclude on the causality between information demand and information supply they performed a Granger Causality test. A granger causality test can yield four different outcomes: a bi-directional causality, unidirectional causality (both directions), and no causality at all. The Granger Causality tests show a bidirectional causality between the two variables. An explanation could be the fact that both variables interact with each other.

Zhang et al. (2013) examined the volatility increases due to the Covid-19 period. To measure investor attention, they used the Chinese equivalent of Google Search Volumes, Baidu. Their results show a strong relationship between investor attention (Baidu index) and abnormal stock returns. In addition, they had to perform a Granger Causality test. The sample of Zhang et al. showed a bi-directional causality, which means that X Granger-causes Y and vice versa. Although, the effect of the abnormal returns to investor attention is less clear. One explanation can be the well-known overconfidence bias by individual investors.

Another study that performed a Granger Causality test is research by Vozlyublennaia, N. (2014). His study focuses on the effect of investor attention on stock index returns. Therefore, the focus of this research was not on individual stocks or companies. In addition, he also used Google Searches as a proxy for investor attention. The results show that investor attention influences the returns of these indexes in the short term. However, the returns influence investor attention in the long run, the opposite direction. The Granger Causality tests show that there is a bi-directional causality.

Table 4. Overview of literature on Granger Causality Tests

Author(s) (Publication year)	Time period	Region	Method	Results
Roberts & Whited (2011)	2012	Worldwide	Descriptive research to endogeneity problems	Endogeneity problems and solutions
Zhang et al. (2013)	2011-2012	China	Correlation analysis and Granger Causality based on Baidu Index	Investor attention has a strong relationship with abnormal return
Vozlyublennaia, N. (2014)	2004-2012	US	VAR model and Granger Causality based on GSV data	Bidirectional relationship investor attention and stock returns.

2.6 Hypotheses

Based on the literature discussed above and what is known to date, expectations can be drawn for the current study. The literature on investor attention in general and investor attention during other crises have led to these expectations. Therefore, several hypotheses have been formulated:

H1: Investor attention Granger causes returns in S&P 500 and S&P 400 mid-cap healthcare companies.

The Granger Causality test is discussed in the literature review and has a crucial role in the preliminary work of this research. In this thesis, investor attention (X) is expected to have an effect on stock returns (Y). Therefore, the expectation is that X Granger-cause Y. In addition, this is also the desired outcome for the reliability of this study.

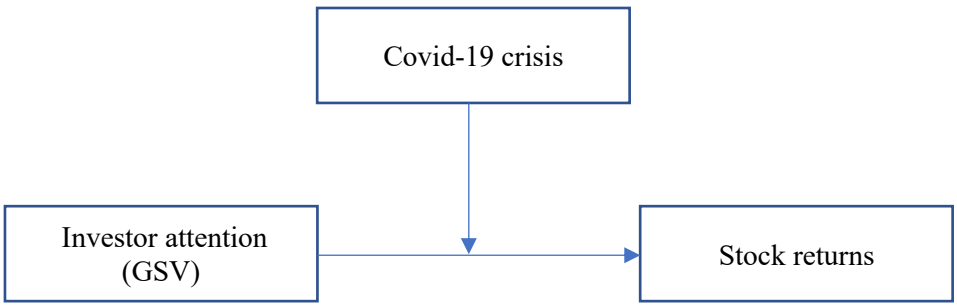
H2a: During Covid-19 (from March 2020), investor attention for healthcare stocks from the S&P 500 exceeds the investor attention in the previous period (2019-Feb 2020).
H2b: During Covid-19 (from March 2020), investor attention for healthcare stocks from the S&P 400 mid-cap exceeds the investor attention in the previous period (2019-Feb 2020).

Healthcare companies are expected to receive more attention during the Covid-19 crisis than in the preceding period. The main reason for this is the health nature of the crisis. In addition, there are many companies that are directly related to the Coronavirus. As a result, these companies have been in the news frequently.

H3a: Investor attention for a healthcare company (large- and mid-cap) has a positive effect on the stock return in the short term (first 1-2 weeks) during the pre-Covid period.
H3b: Investor attention for a healthcare company (large- and mid-cap) has a negative effect on the stock return in the long term (after reversal point) during the pre-Covid period.

A relationship between the concept of investor attention, measured by GSV or Baidu, and stock performance (prices and volatility) has been established by almost every study on this topic. (Bank et al., 2011; DA et al., 2011; Preis et al., 2013; Ying et al., 2015; Bijl et al., 2016). However, there is no unambiguous conclusion about the direction of the effect on stock prices. In addition, Kim et al. (2019) does not find a correlation between Google Searches and stock returns. Nevertheless, it is expected that there is both a positive and negative effect of investor attention for a healthcare company on the stock returns, depending on the short or long term. Most studies found a positive effect in the short term, but a reversal occurred after a certain time market (Bank et al., 2011; DA et al., 2011; Preis et al., 2013; Ying et al., 2015; Bijl et al., 2016). For this reason, these are the expectations for this healthcare stock research regarding the pre-Covid period.

H4a: The effect of investor attention for a healthcare company (large- and mid-cap) on the stock return is more positive in the short term (first 1-2 weeks) during the Covid-19 crisis compared to the pre-Covid period.
H4b: The effect of investor attention for a healthcare company (large- and mid-cap) on the stock return is more positive in the long term (after reversal point) during the Covid-19 crisis compared to the pre-Covid period.



Based on previous research, it can be concluded that a relationship has been established between Google Searches and stock performance. Yu & Hsieh (2010) concluded that the effect of investor attention decreases during the financial crisis, because investors tend to invest less emotionally. Despite their conclusion about the financial crisis, a more positive effect on healthcare stocks is expected during Covid-19 compared to the period before Covid-19. This is mainly due to the health nature of the Covid-19 crisis. It is assumed that healthcare companies received more positive attention during this crisis, as the company names are frequently mentioned on the news and other media sources in a positive way. Smales (2020) found that healthcare stock returns are positively related to searches as 'coronavirus' and 'Covid-19'. Since these companies are fighting the virus, and with it the crisis, this will generate positive attention. In this it can differ from companies from other sectors, which are more likely to experience a decrease in the effect of GSV.

H5a: During Covid-19 (from March 2020), healthcare companies directly related to Covid-19 receive more attention than healthcare companies not directly related to Covid-19.

The fifth hypothesis predicts a distinction between companies directly related to the Covid-19 crisis and companies not directly related to the virus. It is expected that companies directly related to Covid-19, such as companies developing a vaccine or Covid test, receive more attention compared to the other healthcare companies. During the Covid-19 crisis these companies receive more attention on the news and from other media channels. Therefore, it is hypothesized that this will lead to more investor attention and a possible effect on the stock price.

H5b: The effect of investor attention for a healthcare company on the stock return is higher for healthcare companies directly related to the Covid-19 crisis than for healthcare companies not directly related to Covid-19 in both the short and long term.

Hypothesis 4 expects a more positive effect of investor attention on healthcare companies during Covid-19. In addition, hypothesis 5a state that companies directly related to Covid-19 receive more attention than other healthcare companies. Hypothesis 5b combines the previous mentioned hypotheses. Therefore, it is expected that the positive effect of both the Covid-19 crisis and the direct

relation to this crisis will lead to a greater effect on the stock price of companies directly related to Covid-19 compared to the other companies.

H5c: The effect of investor attention for directly related healthcare companies on the stock return is higher during Covid-19 than in the pre-Covid period in both the short and long term.

Since companies directly related to Covid-19 only 'benefit' from this relatedness after the outbreak, it is expected that the effect of investor attention for the directly related companies will be more positive in the period after the outbreak than in the period before.

H6a: Large healthcare companies (based on firm size) directly related to Covid-19 receive the most investor attention during Covid-19.
H6b: The effect of investor attention on the stock returns is the highest for directly related companies with the largest firm size.

It is expected that firm size and relatedness to Covid-19 contribute to the amount of investor attention a company receives. As mentioned before, firm size seems to be positively related to investor attention based on previous literature. Therefore, it is hypothesized that large firms that are directly related to Covid-19 will receive the most attention.

H7a: The effect of investor attention on stock returns for healthcare companies differs across healthcare subsectors.
H7b: The effect of investor attention on stock returns for healthcare subsectors are more positive during Covid-19 compared to the period before.

Given the diversity of the healthcare sector, it is expected that different subsectors are affected in different ways. For example, subsectors including many firms that are related to Covid-19 will receive more attention and the effect will be different. In addition, it is hypothesized that the individual effects per subsector will be more positive in the period after the Covid-19 outbreak, regardless of the subsector. This is a continuation of the previously formulated hypotheses.

3. Data

The data that is used in this research is obtained from Google Trends and WRDS. First of all, the Google Search Volumes (GSV) will be based on Google Trend data of company tickers and company names. In addition, information about the stock prices and other company information will be obtained from the CRSP database. The selection of companies is based on the S&P 500, which consists of the biggest 500 companies from the US stock market. From this index 66 Healthcare companies have been selected that will be examined in the current study (Appendix A1). One of the main reasons that a selection of healthcare stocks from the S&P 500 was chosen is the fact that these stocks are likely to contain the most GSV data. In addition, the healthcare firms that are part of the S&P 400 mid-cap are selected as well. This index contains a total of 36 healthcare firms (Appendix A2). Therefore, the total amount of firms participating in this study is 101. Investor attention can partly be determined by the size and reputation of companies. The S&P 500 contains the largest companies in the world, based on stock market value. Therefore, little distinction can be made in terms of stock market value of the companies and its effect. For this reason, it was decided to also include the S&P 400 mid-cap companies. This makes it possible to distinguish between the largest, best-known companies and the mid-cap companies and possibly clarify the effect of this on the attention that a company receives.

This research aims to examine the effect of the Covid-19 crisis on the effect of investor attention on stock performance. Therefore, a time period of over 2 years will be examined. First of all, the pre-Covid period will consist of the stock performance in 2019 and the start of 2020. In addition, the remaining months of 2020 (March-December) and first months of 2021 (until start of May) form the period after the Covid-19 outbreak. The stock market was extremely volatile the first weeks after the Covid-19 outbreak in mid-March 2020 (Demers et al., 2020). Therefore, 2020 provides an opportunity to examine the effects of investor attention during such a shock. However, a later time period during the Covid-19 period is also included, as this volatile market can lead to distorted results. For this reason, the first months of the year the vaccines were actually developed, 2021, will also be examined in this research.

3.1 Google Trends

Google Trends contains information about the number of searches performed for certain terms. This method, Google Search Volumes, is widely used in the existing literature to measure investor attention. This study uses a similar method regarding Google Searches as Da et al. (2011) and Bijl et al. (2016). In this way information about the number of searches for certain companies and stocks can be determined. This research uses the company ticker, a short code that identifies companies, stocks, or indices and company names. The reason company ticker is sometimes preferred over company name is based on research by Yoshinaga and Rocco (2020). According to them, this excludes people who simply want to gather information about a company and have no intentions as an investor. However, the existing literature shows conflicting arguments about the use of company name and

company ticker. For this reason, both will be examined and the results from company name and company ticker count as robustness check on each other.

In order to research the relationship between investor attention and stock performance as accurately as possible, it was decided to use weekly data. Google Trends data frequency decreases if the sample period increases. Therefore, daily data is only available for a sample period of up to 9 months. Partly for this reason, it is decided to use weekly data instead of daily data. In this way the sample period can be extended. In addition, many similar studies on the effect of investor attention on stock returns have also chosen to use weekly GSV data, such as Da et al. (2011) and Bijl et al. (2016). However, Google Trends offers the possibility to select the searches from a specific country or from all over the world. Since this research concerns the US stock market, it was decided to select searches originating from the US. Search results for US stocks are probably best reflected by searches from the US. In addition, it reduces the chance that a company name or ticker has a different meaning in other countries, which can negatively influence the results.

In addition, the data from Google Trends is based on a scale of 0-100, to indicate the number of searches per unit of time. However, the level of a score depends on the other scores of a specific company. For this reason, this score will have to be standardized which ensures that different searches (companies) can be compared with each other. To calculate the standardized GSV (SGSV), the mean is subtracted from the GSV. *n* is the number of observations. The result is then divided by the total sample standard deviation of the time series (σ_{GSV}), As shown in the equation below:

$$SGSV_t = \frac{GSV_{i,t} - \frac{1}{n}\sum_{i=1}^{n} GSV_i}{\sigma_{GSV}}$$

To create more comprehensibility around the equation of the standardized GSV, a number of features is discussed as described by Swamy et al. (2019):

i) De SGSV is created by the searches for a certain keyword compared to searches for all keywords during the same period.
ii) The GSV is a relative value which takes a value between 0-100.
iii) The SGSV shows the relative frequency of searches.
iv) The GSV and its value does not increase due to lower searches for a particular query relative to others.

3.2 CRSP/Compustat

The stock price data is obtained from the CRSP database from WRDS. The selection of companies and stocks is based on the S&P 500, which contains 66 healthcare stocks. In addition, 36 companies

from the S&P 400 mid-cap participate in this thesis. For these companies, the historical daily stock prices are obtained for a time period of three years. First of all, the company name, company ticker, and Cusip are selected as identifying variables. Based on the identifying variables, the dataset can be combined with Google Trends data. Secondly, for the time series information, both the daily stock price and daily return without dividends are added to the dataset of the current research.

In addition, the Compustat database is used to calculate several control variables. First of all, total assets (log) are used to measure firm size. Secondly, Tobin's Q is calculated as a measure to control for replacement value of the assets. Tobin's Q is calculated based on several variables from the Compustat database. In addition to total assets, the shares outstanding, close price, and common equity (book value) are also included to calculate this. Other control variables that are included in this study are return on equity, leverage, and the market-to-book ratio. In general, intangible assets cover a large part of the total assets compared to other sectors. Therefore, goodwill and total intangible assets are also controlled for. The relation between intangible assets and stock- or equity returns is studied before (Doherty et al., 2011). In addition, Kedron (2020) showed the relationship between goodwill and share prices. Unfortunately, research and development costs are excluded from this research due to a lack of available data. In addition, there was no data on the variable 'trading volume' for every company, therefore this variable was also excluded from the study. Finally, there will be controlled for inflation based on the CPI index from the Bureau of Labor Statistics.

Due to a lack of complete stock data, some of the 66 healthcare companies in the S&P 500 were excluded from this study. Viatris Inc. (VTRS) did not enter the stock market until November 2020 after emerging from Pfizer Inc. and Mylan N.V. In addition, the shares of Organon & Co (OGN) were first traded on the stock exchange in May 2021. This took place after a spin-off from Merck & Co. (MRK). BT Group PLC. (BT) was excluded from this study because stock data was only available on CRSP until September 2019. Therefore, this thesis studies 62 healthcare companies of the S&P 500. In addition, several healthcare companies from the S&P 400 mid-cap are excluded. Inari Medical (NARI), Envista Holdings (NVST), Option Care Health (OPCH) and Progyny (PGNY) are excluded from this study, as part of the stock data during the sample period is not complete. As a result, 32 of the 36 healthcare companies remain from the S&P 400 mid-cap. Therefore, this study includes a total of 95 companies.

In this section the subsectors will be discussed based on a classification of 9 subsectors. Table 5 contains information regarding these subsectors and corresponding SIC codes. In addition, it indicates how many companies from the healthcare sectors of the S&P 500 and S&P 400 mid-cap belong to the relevant subsector. The first and bold SIC codes are the most common SIC code for this subsector. A small explanation will be given in column 3. These subsectors will be used in the research on hypothesis 7.

Table 5. Subsectors and SIC codes

Table 5 shows the 9 different subsector and corresponding SIC codes. In addition, a short explanation can be found in column 3. The number of firms that belongs to a specific subsector can be found in column 4, based on firms that participate to this study.

Subsector	***SIC code***	***Explanation***	***# Firms***
Biotechnology	***2836,*** *2834, 8731*	*Codes starting with 2834 belongs to 'Drugs'. SIC code 2836 includes especially biological products.*	*13*
Healthcare Distribution	***5122, 5047,*** *3826, 2834*	*5122 mostly engaged in distribution of drugs. 5047 concerns distribution of hospital, dental and surgical equipment.*	*7*
Healthcare Equipment	***3841, 3845,*** *3826, 3842, 3844, 3851, 2834, 2835, 8090*	*SIC code 384 contains medical instruments and supplies. 3841 (surgical and medical instruments) and 3845 (electromedical equipment) are the most common in this subsector.*	*28*
Healthcare facilities	***8062,*** *8060, 8090,*	*SIC code 806 is called hospitals. 8062 (general medical and surgical hospitals) and 8060 (services hospitals) are the most common in this subsector.*	*6*
Healthcare Services	***8071, 8082,*** *8000, 8741, 7373*	*SIC code 80 includes Health Services. 8071 (medical laboratories) and 8082 (home healthcare services) are the most common in this subsector.*	*8*
Healthcare Supplies	***3841, 3843,*** *3851, 3060 2834, 2835,*	*SIC code 384 includes medical instruments and supplies. 3841 (surgical and medical instruments) and 3843 (Dental equipment and supplies) are the most common in this subsector.*	*8*
Lifesciences Tools & Services	***8731, 3826,*** *2836,*	*8731 contains commercial, physical and biological research. 3826 includes laboratory analytical instruments.*	*9*
Managed Healthcare	***6324,*** *7370*	*6324 is called hospital and medical service plans.*	*7*
Pharmaceuticals	***2834, 2836,***	*2834 is pharmaceutical preparations and 2836 are biological products.*	*9*

3.3 Control variables

To conduct this research, some control variables will be added. First of all, there will be controlled for subsector-, firm- and time-fixed effect, due to the use of panel data with companies in different subsectors. Since weekly data is examined over a longer period of time, time-fixed effects are also

controlled for. Secondly, several control variables are added to the study. This first variable is firm size, which is correlated to investor attention according to Sabbaghi (2022). Firm size is measured by the log of the total assets of a company. In addition, there will be controlled for the replacement of assets (Tobin's Q). This is calculated by multiplying the shares outstanding by the fiscal close price. The total assets are added to the outcome while the common equity is subtracted. Hereafter, the total amount of the calculation above is divided by the total assets. In addition, the return on equity (RoE) is calculated by dividing the net income by total equity. Next, the leverage of a firm is measured by dividing the total debt by the shareholders equity. To counteract the extreme values in terms of leverage, it has been decided to winsorize the data at 1%. Due to winsorization, the extreme outliers are replaced by the next largest/smallest value. Therefore, the results are easier to interpret. Finally, the market to book ratio is calculated by the market value divided by the book value per share times the shares outstanding. The outcomes of the market to book ratio are also winsorized at 1% to avoid extreme outliers that can influence the regressions. In addition, the logarithms of goodwill, intangible assets, and capital expenditures have also been taken, in order to make the large differences easier to interpret. The descriptive statistics including the adjusted control variables can be found in Appendix B.

3.4 Descriptive statistics

In this section the descriptive statistics will be discussed. These descriptive statistics can provide a first insight into the variables and the relationship between certain variables. Below in Table 6 to Table 10, the descriptive statistics related to all variables can be found. This concerns the number of observations, the mean, the median, the minimum, the maximum, and the standard deviation. In addition, a distinction is made between the pre-Covid period and the period during Covid-19 for the dependent and independent variables. For the GSV data, all the aforementioned measures can be found for the total number of observations, but also for the pre- and post-covid-19 outbreak separately. The same principle applies to the data regarding the weekly stock returns. Finally, the descriptive statistics about the control variables can be found at the bottom.

Table 6. Descriptive Statistics SGSV by Ticker

Table 6 includes the raw descriptive statistics of the SGSV data measured by Ticker. Therefore, the number of observations, mean, median, minimum, maximum, and standard deviation are shown.

	N	Mean	Median	Min	Max	St. Dev
SGSV pre	5795	-.045	-.192	-.2629	7.301	.972
SGSV post	5795	.070	-.081	-.3.299	10.249	1.009
SGSV Tot.	11590	.017	-.145	-3.299	10.249	.991

Table 7. Descriptive Statistics SGSV by Company Name

Table 7 includes the raw descriptive statistics of the SGSV data measured by Company Name. Therefore, the number of observations, mean, median, minimum, maximum, and standard deviation are shown.

	N	Mean	Median	Min	Max	St. Dev
SGSV pre	5795	.082	-.021	-3.926	8.375	.977
SGSV post	5795	-.030	-.195	-3.378	12.003	1.042
SGSV Tot.	11590	.006	-.117	-3.926	12.003	1.013

Tables 6 and 7 show the descriptive statistics of the Google Search Volume (GSV) data. As mentioned in the previous section, the GSV data is standardized to a Standardized Google Search Volume (SGSV). The descriptive statistics are based on these standardized GSV (SGSV) data. It is expected that the companies from the sample receive more attention due to Covid-19. Table 6 shows statistics that match this expectation. The mean of the SGSV data in the period after the Covid-19 outbreak is the highest. This means that the healthcare companies from the S&P 500 and S&P 400 mid-cap received more attention during this period, based on search volumes. In addition, the mean of the period before Covid-19 is negative. Therefore, it can be concluded that these results correspond to the expectation that Covid-19 will cause an increase in investor attention for healthcare companies from the US. However, Table 7 shows partly conflicting statistics. The SGSV data based on company name shows that the healthcare companies from the S&P 500 and 400 mid-cap attracted more attention during the period before Covid-19, compared to the other period. In this case, the descriptive statistics are less in line with the expectations.

There are several reasons for deviations in the accuracy of the GSV results for both the company name and the company ticker. Therefore, various choices are made in the existing literature that conducts similar studies. For example, a company name can also be searched for by people who are interested in the company or product, and not as an investor. In addition, a company ticker can also have other meanings, especially if the searches are included worldwide. For this reason, it is interesting to take a combination of both search results. In Table 8 these pooled results can be found.

To conclude on hypothesis 6, a distinction will be made between firm sizes based on the S&P 500 and 400 mid-cap. A potential reason for the conflicting descriptive statistics could be the fact that smaller companies received less attention due to Covid-19. To answer hypothesis 6, separate analysis will be performed.

Table 8. Descriptive Statistics SGSV Combined

Table 8 includes the raw descriptive statistics of the SGSV data measured by the combination of Ticker and Company Name searches. Therefore, the number of observations, mean, median, minimum, maximum, and standard deviation are shown.

	N	Mean	Median	Min	Max	St. Dev
SGSV pre	5795	.019	-.053	-2.848	6.583	.784
SGSV post	5795	.020	-.082	-2.858	9.808	.839
SGSV Tot.	11590	.012	-.077	-2.858	9.808	.812

Table 8 shows the combined results from Table 6 and 7. The results are in line with expectations that the attention, measured on the basis of Google search results, for healthcare companies in the S&P 500 and 400 mid-cap will increase due to the Covid-19 crisis. Although, the increase is nihil (.020 > .019). Therefore, it does not completely match the expectation. A possible explanation could be the fact that firms from the S&P 400 mid-cap, which are also included in these results, receive less attention during Covid-19, relatively. In addition, as in Tables 6 and 7, the period during Covid-19 shows the largest standard deviation and the largest difference between the minimum and maximum. During times of crisis, a higher volatility is expected.

Table 9. Descriptive Statistics Stock Returns

Table 9 includes the raw descriptive statistics of the stock returns. Therefore, the number of observations, mean, median, minimum, maximum, and standard deviation are shown.

	N	Mean	Median	Min	Max	St. Dev
Return pre	5795	.0007	.0011	-.0731	.1059	.0099
Return post	5795	.0019	.0015	-.0912	.1426	.0134
Return Tot.	11590	.0013	.0011	-.0912	.1426	.012

In Table 9 the descriptive statistics of the dependent variable, the stock returns, can be found. These weekly stock returns are calculated by taking the weekly average based on daily returns. Since the stock market is closed during the weekend (Saturday and Sunday), there are a maximum of five returns per week. However, there were specific weeks in which the number of days the exchange was open was only three or four. The statistics show small differences between the years. Compared to the pre-Covid period, the mean almost doubled in the period during Covid-19. Therefore, it can be concluded that the stock returns were slightly higher after the outbreak. In addition, the standard deviation and the difference between the minimum and maximum are also the during Covid-19.

Table 10. Descriptive Statistics Control Variables

Table 10 includes the raw descriptive statistics of the control variables. Therefore, the number of observations, mean, median, minimum, maximum, and standard deviation are shown per year.

	N	Mean	Median	Min	Max	St. Dev
Tobin's Q 2019	5035	3.578	2.654	1.059	13.083	2.466
Tobin's Q 2020	4940	3.998	2.883	1.026	19.340	3.158
Tobin's Q 2021	1615	4.155	2.594	1.005	23.563	3.569
Firm size 2019	5035	3.894	3.832	2.316	5.347	.711
Firm size 2020	4940	3.966	3.880	2.411	5.363	.684
Firm size 2021	1615	4.028	3.983	2.539	5.367	.668
RoE 2019	5035	0.324	.026	-.078	.159	.033
RoE 2020	4940	.017	.023	-.336	.219	.068
RoE 2021	1615	.038	.030	-.042	.188	.037
Leverage 2019	5035	.146	.614	-42.501	6.062	4.926
Leverage 2020	4940	3.223	.634	-447.857	603.786	77.319
Leverage 2021	1615	1.175	.617	-39.363	34.946	6.020
MTB 2019	5035	6.057	4.546	-67.671	125.761	17.956
MTB 2020	4940	-3.479	5.709	-1467.57	151.25	152.706
MTB 2021	1615	10.865	5.073	-84.120	226.165	28.719
Goodwill 2019	5035	7410.597	2220.9	0	79749	14074.36
Goodwill 2020	4940	8034.171	2105.264	0	79552	14572.18
Goodwill 2021	1615	8719.925	2504.202	0	79121	15097.7
Intangible assets 2019	5035	12279.03	2781.5	0	112870	22957.86
Intangible assets 2020	4940	13725.97	2750	0	116000	25125.78
Intangible assets 2021	1615	14413.55	3587.125	0	108330	24603.19
CPI 2019	5035	255.650	255.9	252.47	258.263	1.709
CPI 2020	4940	258.8446	258.845	255.944	261.564	1.677
CPI 2021	1615	264.367	265.028	262.2	266.727	1.671
Cap ex 2019	5035	474.031	155	4.04	4158	752.999
Cap ex 2020	4940	495.945	187.379	2.504	4684	772.137
Cap ex 2021	1615	561.996	244.811	1.457	4448	841.134

In Table 10 the descriptive statistics regarding the control variables can be found. Given that the control variables are available per year, the descriptive statistics are also presented per year. Companies always strive for growth and improvement and this is reflected in the figures per year. Both Tobin's Q and Firm Size increase slightly each year. The same phenomenon yields for the inflation per year, measured by CPI. Remarkable is the fact that the mean of the ratio based on net income (RoE) is the lowest in 2020 to compared to 2019 and 2021. During 2020, companies have generally been hit the hardest which is reflected in this. In addition, the high volatility of almost all control variables is noticeable, expressed in the standard deviation and the size difference in the minimum and maximum. Previous studies have already shown that higher investor attention and a period of crisis are associated with high volatility. The high volatility in 2020 is also reflected in the market-to-book ratio and leverage, since the standard deviation for the market to book ratio and leverage is the highest during 2020 compared to the other years.

3.5 Vector Autoregression Model (VAR)

In this section the Vector Autoregression Model will be discussed, followed by an explanation of the Granger Causality test. The focus in a Vector Autoregression Model is based on two variables, the independent and dependent variable. In this thesis the dependent variable are the stock returns and the independent variable consist of the GSV data. To perform a Granger Causality test, a Vector Autoregression model is needed. The results of the Granger Causality test can be used to answer the hypotheses regarding this test. Below the Vector Autoregression Model will be explained.

A Vector Autoregression Model can be used for multivariate time series models. A multivariate time series model is a dataset consisting of more than a single time-dependent variable. Time-dependent variables vary over time. The idea behind the VAR model is that it uses the past values within the time series as a predictor for current or future values. This research will examine whether investor attention can predict stock returns for healthcare companies within the S&P 500. At the same time, it will be analyzed whether stock returns can predict investor attention for the same group of healthcare companies. The explanatory variables are the lagged values of the concerning variables. In other words, a VAR model regresses a vector of time series variables on its own lags. In addition, the coefficients can be estimated by using Ordinary Least Squares (OLS) for each equation. Finally, the equations that are used for the VAR model in this thesis are similar to the two equations in the next section on the Granger Causality tests.

3.6 Granger Causality tests

A Granger Causality test will be performed to conclude on the causality between the dependent and independent variables. It is important to perform this test if time series or panel data are used. The Granger Causality test uses lag variables of one variable (X) to predict future values of the other

variable (Y). In addition, the test is also performed in the opposite direction, from the lags of Y to predicting the future values of X. The objective of the Granger Causality test is to conclude on which variable comes before the other variable in time series data. In this thesis, it is expected that investor attention will have an effect on stock returns. Although, it could be the case that high stock returns will lead to more investor attention. If the conclusion is that variable X Granger causes variable Y, it means that lag variables of X (past data) predict future values of variable Y. To perform the Granger Causality test, this will be tested in both directions.

The equations that are used for the Granger Causality test are similar to the equations of the Vector Autoregression Model (VAR). Since the Granger Causality test will be performed in both directions, there are two equations. The following two equations are used to conclude on the causality between investor attention (measured by GSV data) and stock returns:

$$\boldsymbol{R_{i,t} = \alpha_i + \sum_{N=1}^{N} \beta_{1N} * SGSV_{i,t-N} + \sum_{N=1}^{N} \beta_{2N} * R_{i,t-N} + y * C_{i,t} + S_i + F_i + T_t + \varepsilon_{i,t}}$$

$$\boldsymbol{SGSV_{i,t} = \alpha_i + \sum_{N=1}^{N} \beta_{1N} * SGSV_{i,t-N} + \sum_{N=1}^{N} \beta_{2N} * R_{i,t-N} + y * C_{i,t} + S_i + F_i + T_t + \varepsilon_{i,t}}$$

Both equations are identical to each other, except from the dependent variable. In the first equation the dependent variable is the weekly stock return $(R_{i,t})$. The second equation uses weekly standardized GSV data $(SGSV_{i,t})$ as the dependent variable. The remaining part of the equation consists of the explanatory variables. These variables are formed by the lagged variables of investor attention $(SGSV_{i,t-N})$ and the stock returns $(R_{i,t-N})$. In addition, the dependent variables are explained by a constant (α_i), the vector of coefficients (y) and the control variables $(C_{i,t})$, which are not lagged. The control variables, consisting of Tobin's Q and Firm Size, are multiplied by this vector of coefficients. Furthermore, the subsector fixed effects $(\boldsymbol{S_i})$, firm fixed effects $(\boldsymbol{F_i})$ and the time fixed effects $(\boldsymbol{T_t})$ have been added to the equation. Finally, $\boldsymbol{\varepsilon_{i,t}}$ reflects the error term of all fixed effects.

The Granger Causality tests can have different outcomes. It can conclude that there is a unidirectional causality between the two investigated variables. In this situation, X Granger-cause Y or Y Granger-cause X. In other words, the lagged values of X or Y can predict the future values of Y or X, respectively. In addition, the other possibility is that the lagged values of X or Y can predict the future values of both X and Y. Then, the outcome of the Granger Causality test will be that there is a bidirectional causality.

3.7 Hausman test

Before running the regressions, it must be determined by means of a Hausman test whether a fixed effects or random effects model will be used. Therefore, the regression (including the control variables) will be run twice. One regression including the fixed effects command, and the other analysis including the random effects command in Stata. The fixed effects that are included are firm, subsector, and week fixed effects. The Hausman test will compare both results and conclude on which effects should be used for the following regressions. This is determined on the basis of a null and alternative hypothesis. The null hypothesis is as follows: 'the random effects model is the preferred model'. In addition, the alternative hypothesis is the following: 'the fixed effects model is the preferred model' The results, based on the regression including SGSV Combined data, can be found in Table 11.

Table 11. Hausman Test firm fixed effects (SGSV Combined)

Table 11 shows the results of the Hausman test based on the SGSV Combined measure, which indicates whether a fixed effects or random effects model is preferred.

Test	Value
Chi Square	43.73
P-value	.0001

The null hypothesis can be rejected, since the P-value is significant at 5% level (.0001 > .05). Therefore, the firm fixed effects model is the preferred model. Therefore, the aforementioned fixed effects will be used for the regressions, both separately and jointly In Appendix C the results of the Hausman test with SGSV by ticker and company name can be found. All SGSV measures give the same result.

4. Methodology

This research will examine the effect of investor attention on stock returns during Covid-19 in the US. Therefore, a comparison between the pre-Covid period (2019-Feb 2020) with the post-Covid outbreak period (March 2020-2021) will be made. The healthcare stocks from the S&P 500 and S&P 400 mid-cap will be examined, due to the health nature of the crisis and availability of data. In this section the methodology will be discussed per hypothesis.

H1: Investor attention Granger-causes stock returns in S&P 500 and S&P 400 mid-cap healthcare companies.

Hypothesis 1 will be answered by performing the previous explained Granger Causality tests. To perform this research, it would be helpful if X (investor attention) causes Y (stock returns). Firstly, the Vector Autoregression Model (VAR) needs to be set up. Hereafter, the Granger Causality tests can be performed by Stata. The following two equations will be used to conclude on the causality between investor attention and stock returns for healthcare companies in the S&P 500 and S&P 400 mid-cap:

$$\boldsymbol{R_{i,t} = \alpha_i + \sum_{N=1}^{N} \beta_{1N} * SGSV_{i,t-N} + \sum_{N=1}^{N} \beta_{2N} * R_{i,t-N} + y * C_{i,t} + S_i + F_i + T_t + \varepsilon_{i,t}}$$

$$\boldsymbol{SGSV_{i,t} = \alpha_i + \sum_{N=1}^{N} \beta_{1N} * SGSV_{i,t-N} + \sum_{N=1}^{N} \beta_{2N} * R_{i,t-N} + y * C_{i,t} + S_i + F_i + T_t + \varepsilon_{i,t}}$$

H2a: During Covid-19 (from March 2020), investor attention for healthcare stocks from the S&P 500 exceeds the investor attention in the previous period (2019-Feb 2020).
H2b: During Covid-19 (from March 2020), investor attention for healthcare stocks from the S&P 400 mid-cap exceeds the investor attention in the previous period (2019-Feb 2020).

The second hypothesis can be answered on the basis of an analysis of the descriptive statistics. Annual data are used to make a comparison between the pre-Covid period (Jan 2019-Feb 2020) and the period after the outbreak of Covid-19 (Mar 2020-Dec 2021). In order to make a comparison, the standardized Google Search Volumes have been calculated. On the basis of the mean per month/year, a comparison will be made between the investor attention in both periods.

H3a: Investor attention for a healthcare company (large- and mid-cap) has a positive effect on the stock return in the short term (first 1-2 weeks) during the pre-Covid period.
H3b: Investor attention for a healthcare company (large- and mid-cap) has a negative effect on the stock return in the long term (after reversal point) during the pre-Covid period.

H4a: The effect of investor attention for a healthcare company (large- and mid-cap) on the stock return is more positive in the short term (first 1-2 weeks) during the Covid-19 crisis compared to the pre-Covid period.
H4b: The effect of investor attention for a healthcare company (large- and mid-cap) on the stock return is more positive in the long term (after reversal point) during the Covid-19 crisis compared to the pre-Covid period.

Since weekly data is used for both stock prices and searches (GSV) for multiple companies, a panel data analysis is required. In addition, a fixed effects regression will be performed to control for industry, firm and week fixed effects. The regression equations for both hypotheses 3 and 4 are the following:

$$\boldsymbol{R_{i,t} = S_i + F_i + T_t + \sum_{L=1}^{L} \beta_{1L} * SGSV_{i,t-L} + \sum_{L=1}^{L} \beta_{2L} * C_{i,t-L} + \varepsilon_{i,t}}$$

$R_{i,t}$ is the dependent variable and shows the stock return per company and unit of time. The degree of investor attention per company and unit of time is measured using standardized Google Search Volumes ($SGSV_{i,t}$). In this study, the SGSV is the variable of interest. The fixed effects are locked into S_i, F_i and T_t in the regression equation. The industry fixed effects are covered by S_i, the firm fixed effects by F_i, and the time fixed effects are covered by T_t. In addition, a number of other variables must be controlled for, based on the existing literature. These variables, discussed in the data section, are represented by $C_{i,t}$. Finally, the L represents the number of lags of the SGSV and control variables that are included. To conclude on the second hypothesis, the results of the post-Covid period will be compared to the results of the pre-Covid period. However, for both periods the same methodology will be used.

Since the existing literature agrees that there is a difference in the effect of GSV in the short and long term (reversal point), the effect on a later time period will be studied as well (Bank et al., 2011; DA et al., 2011; Preis et al., 2013; Ying et al., 2015; Bijl et al., 2016). For this, lag variables of the SGSV measures will be used. However, these studies do not fully agree on when this reversal point occurs. Hence, several lag variables are included.

H5a: During Covid-19 (from March 2020), healthcare companies directly related to Covid-19 receive more attention than healthcare companies not directly related to Covid-19.

In order to draw a conclusion about hypothesis 5, a distinction is made between healthcare companies directly related to Covid-19 and healthcare companies not directly related to the virus or the crisis.

Companies directly related to Covid-19 can be related to the virus in a variety of ways. An example are companies that participate in the development, distribution or issuing of a vaccine. Another possibility are companies that are specialized in respiratory equipment and that have specifically committed themselves to corona patients. In order to conclude about the difference in the amount of investor attention between the two groups of companies, an analysis of the descriptive statistics will be performed for each group of companies. A list of all companies that are considered as directly related to Covid-19 can be found in Appendix D.

H5b: The effect of investor attention for a healthcare company on the stock return is higher for healthcare companies directly related to the Covid-19 crisis than for healthcare companies not directly related to Covid-19 in both the short and long term.

To conclude on the fifth hypothesis, a similar method will be used as for the second and third hypotheses. However, no comparison is made between the pre-Covid period and the period during Covid-19, but between Covid-19 related healthcare companies and unrelated healthcare companies.

H5c: The effect of investor attention for directly related healthcare companies on the stock return is higher during Covid-19 than in the pre-Covid period in both the short and long term.

Hypothesis 5c focusses on the difference in effect of investor attention between the pre- and post-Covid outbreak periods for directly related companies. Therefore, two separate regressions will be performed and compared for the related companies. A similar panel regression will be performed as for the other regression analysis.

H6a: Large healthcare companies (based on firm size) directly related to Covid-19 receive the most investor attention during Covid-19.

Hypothesis 6a is based on a combination of the hypotheses regarding the firm size and relatedness of the healthcare companies. To conclude on this hypothesis, the healthcare companies have been divided into 4 quantiles based on firm size. In addition, the dummy variable that indicates whether a company is directly related to Covid-19 or not, is also used. Therefore, 8 different combinations will be compared to each other. The expectation is that the group that contains related companies with the largest firm sizes, will receive the most attention.

H6b: The effect of investor attention on the stock returns is the highest for directly related companies with the largest firm size.

First, the companies will be divided into 4 groups instead of 8. Quantile 1 and 2 (smallest firm size) and 3 and 4 (largest firm size) will be merged. In addition, the distinction between related and unrelated remains intact. This leaves 4 combinations for which the same regression analysis will be performed as described earlier.

H7a: The effect of investor attention on stock returns differs across the 9 different healthcare subsectors.

In order to draw a conclusion about the effect of investor attention of the individual healthcare subsectors on stock returns, a pooled OLS regression will be performed. The pooled OLS regression is a basic method to analyze panel data. First, a dummy variable is created for each subsector. All dummy variables are included in the pooled OLS regression to analyze the coefficient. Instead of a fixed effects panel regressions, an OLS regression will be performed using panel data. The formula of the regression can be found below:

$$\boldsymbol{R_{i,t}} = \sum_{\boldsymbol{L=1}}^{\boldsymbol{L}} \boldsymbol{\beta_{1L}} * \boldsymbol{SGSV_{i,t-L}} + \sum_{\boldsymbol{L=1}}^{\boldsymbol{L}} \boldsymbol{\beta_{2L}} * \boldsymbol{C_{i,t-L}} + \boldsymbol{\varepsilon_{i,t}}$$

The regression equation does not contain any fixed effects. In addition, the dummy variables for the individual subsectors will be added, except from one. This allows the individual effects per subsector to be examined.

H7b: The effect of investor attention on stock returns for healthcare subsectors are more positive during Covid-19 compared to the period before.

Finally, a comparison will be made between the pooled OLS regressions of the pre- and post-Covid-19 outbreak periods. For this, the same regression is performed as described above. However, two regressions are performed for separate periods.

5. Results

In this section the results will be discussed for each hypothesis separately. First of all, the results regarding the causality between investor attention and stock returns will be discussed (hypothesis 1). Thereafter, the difference regarding the amount of investor attention between the pre-Covid period and the period after the outbreak will be analyzed (hypothesis 2). Furthermore, the effect of investor attention during these two periods will also be discussed. Hypothesis 3 and 4 made a distinction between both the period before and after Covid-19 outbreak and between the short (first 1-2 weeks) and long term (after reversal point). The results of hypothesis 5 will conclude on the differences between companies directly related to Covid-19 and companies not directly related to the virus. In addition, the results of hypothesis 6 will show the differences of the effect of investor attention between companies from the S&P 500 and S&P 400 mid-cap. Finally, hypothesis 7 will combine the analysis of hypothesis 5 and 6.

5.1 Granger Causality tests results

First, the causality between investor attention and stock returns must be established. Therefore, the Granger Causality tests have to be performed. The first step is to find out how many lags is the optimal number. This can be determined by means of the function 'pvarsoc' in Stata. The function pvarsoc will be used for the SGSV by ticker, SGSV by company name and the combined SGSV. This study focuses on the combined SGSV results. The other two forms of SGSV data form a robustness check on this.

Since the Google Trends data is linked to the stock data from at least one week later, the first layer of the SGSV data is used both for determining the optimal number of lags and for the Granger Causality itself. In other words, the first lag is equivalent to the 'normal' SGSV variable in this study. Below in Table 12, the results can be found regarding the lag determination. Appendix E contains information on the optimal lag determination for SGSV by ticker and company name.

Table 12. Optimal Lag Determination

Table 12 includes the results to determine the optimal number of lags to perform the Granger Causality tests, based on the SGSV Combined measure. The lowest values of MBIC, MAIC, and MQIC determine the number of lags jointly. In addition, these lowest values are indicated by *.

Lag	CD	J	J Value	MBIC	MAIC	MQIC
1	-.122	90.073	.000	-62.636	58.073	-17.882
2	-1.337	42.239	.000	-72.293*	18.239	-11.904
3	-2.424	21.724	.005	-54.630	5.724	-14.371
4	-3.900	2.300	.681	-35.878	-5.701*	-15.748*

Based on the results in Table 12, it can be determined that the optimal number of lags for the Granger Causality test is 4 lags. To arrive at this result, it is necessary to determine the lowest outcomes among MBIC, MAIC and MQIC. In Table 12 these lowest values are indicated by *. In this specific case, it concerns lag 2 (MBIC) and two times lag 4 (MAIC and MQIC). Therefore, it can be concluded that a number of 4 lags is optimal. Appendix E includes the results of the other SGSV measures regarding the optimal lag determination and corresponding Granger Causality tests.

The next step is to perform the Granger Causality tests. For this test the number of 4 lags will be used, as determined in Table 12. The outcomes of the Granger Causality test can be found in Table 13 below. Moreover, the same steps will also be performed for SGSV by Ticker and SGSV by company name. The results of these tests will be used as robustness checks on the results discussed in this section and can be found in Appendix E.

Table 13. Granger Causality tests (SGSV Combined)

Table 13 shows the results of the Granger Causality tests based on the SGSV Combined measure. Therefore, a Chi-square test is performed. The direction of the causality between investor attention (SGSV) and stock returns (Weekly Average) can be determined. ***, **, and *, indicate the 1%, 5%, and 10% significance levels, respectively.

Dependent var.	Excluded	Chi-sq	df	p-value
Weekly Average	SGSV Combined	38.304	4	.000***
	ALL	38.304	4	.000***
SGSV Combined	Weekly Average	10.519	4	.033**
	All	10.519	4	.033**

Table 13 shows the results regarding the causality between the Google Search data (SGSV Combined) and the stock returns (Weekly Average). The null hypothesis of the first part of the Granger Causality test is the following: Investor attention does not Granger-cause stock returns for healthcare companies from S&P 500 and S&P 400 mid-cap. This null hypothesis can be rejected at 1% significance level, given the small p-value (.000). For the second part of the test, the null hypothesis is as follows: Stock returns does not Granger-cause investor attention for healthcare companies from S&P 500 and S&P 400 mid-cap. The p-value of .033 shows that the null hypothesis can be rejected at 5% significance level. Therefore, a bidirectional causality can be established. Although, at the 1% significance level only a unidirectional causality can be derived (investor attention Granger-cause stock returns).

5.2 Optimal lag determination panel regressions

Before the fixed effects regressions can be performed with panel data, the optimal number of lags per explanatory variable must be determined. Therefore, a normal panel data regression is first performed with a total of 6 lags of each variable. Based on the significance, it can be determined to which lag the

SGSV data plays an explanatory role. The results can be found in Appendix F. This shows that the SGSV data has a significant effect for week 1 to week 4. Week 5 and 6 do not show any significance. Therefore, it can be concluded that the optimal number of lags is 4 weeks. In addition, Appendix F shows the results for SGSV by ticker and by company name as robustness checks.

5.3 Amount of investor attention before and during Covid-19

Hypothesis 2 made two distinctions. First of all, companies from the S&P 500 will be analyzed separately from companies from the S&P 400 mid-cap. In addition, the period before the Covid-19 outbreak and the period after this outbreak will be compared for both subsets of companies. To conclude on hypothesis 2, an analysis of several descriptive statistics will be performed. Table 14 shows the amount of investor attention for both groups of companies during the pre-Covid period and the period during Covid-19. The pre-Covid period includes 2019 and the first months of 2020. In addition, the period during Covid-19 contains the remaining months of 2020 and 2021.

Table 14. SGSVs pre- and during Covid-19

The amount of investor attention, based on all three measures, are shown in Table 14. Therefore, the results are divided into four groups, based on S&P500/S&P 400 and pre-Covid/during Covid.

	SGSV by Ticker	SGSV by Company Name	SGSV Combined
Pre Covid & S&P 500	-.114	.029	-.042
During Covid & S&P 500	.134	-.055	.040
Pre-Covid & S&P 400 mid-cap	.146	.187	.166
During Covid & S&P 400 mid-cap	-.083	-.010	-.092

Table 14 indicates conflicting results regarding the amount of investor attention based on different search methods. The expectation was that both companies from the S&P 500 and S&P 400 mid-cap will exceeds the amount of investor attention in the period during Covid-19 compared to the period before. Hypothesis 2a can be partly confirmed (S&P 500 companies), as two out of three measures show a higher investor attention during Covid-19 (SGSV by Ticker and Combined). However, the results for the S&P 400 mid-cap companies are completely contradictory to the expectations and hypothesis 2b. According to the descriptive statistics in Table 14, the amount of investor attention for S&P 400 companies will only decrease during Covid-19. A potential explanation for this can be the fact that the S&P 500 contain different kind of companies. The list of directly related companies is

covered by almost only companies from the S&P 500. Therefore, it can be stated that these companies benefit more from the Covid-19 outbreak in terms of attention.

5.4 Effect of investor attention before Covid-19

For the analysis of hypothesis 3a and 3b, all companies are considered as one group. However, the periods before and during Covid-19 will be compared. To conclude on hypothesis 3, a fixed effects regression analysis will be performed regarding the effect of investor attention on stock returns. There will be controlled for industry, firm and week fixed effects, both separately and jointly. The industry fixed effects are based on the standard classification industry codes (SIC), which are subcategories. Based on the literature review, a positive effect of investor attention on stock returns is expected for the first two weeks. After 3 weeks a price reversal is expected. Table 15 shows the regression results of the fixed effects regression. The regression analysis includes 4 lags (weeks) and is performed for each measure of SGSV data separately. The results for the SGSV by ticker and SGSV by company name measures can be found in Appendix G**.** Therefore, the coefficient, standard error and significance will be showed for SGSV by Ticker, SGSV by Company Name, SGSV Combined and the control variables. The SGSV data by Ticker and Company Name act as a robustness check on the SGVS Combined data. Table 15(1) and 15(2) include 6 different regressions: OLS, Subsector FE, Firm FE, Subsector + Week FE, Firm + Week FE, and Subsector + Firm + Week FE.

Table 15(1). Effect of investor attention by SGSV Combined (pre-Covid)

This table presents three regressions regarding the effect of investor attention during the pre-Covid period, measured by SGSV Combined data. Therefore, the sample period is January 2019 to February 2020. The regressions showed in this table are an OLS, subsector fixed effects, and firm fixed effects regression. Moreover, the table includes 4 lags of the SGSV data, a lag of the dependent variable, and a lag of the control variables. The F-test indicates the significance of the entire model and the R2-adjusted is, compared to the R2, adjusted for the number of predictor variables in the model. ***, **, and *, indicate the 1%, 5%, and 10% significance levels, respectively.

Weekly Average	**OLS**		**SubSector FE**		**Firm FE**	
	Coefficient	**St. Error**	**Coefficient**	**St. Error**	**Coefficient**	**St. Error**
WA_L1	-.0709***	.0144	-.0711***	.0144	-.0788***	.0145
Lag1	.0003*	.0002	.0003*	.0002	.0004*	.0002
Lag2	.0001	.0002	.0001	.0002	.0001	.0002
Lag3	-.0005**	.0002	-.0005**	.0002	-.0005**	.0002
Lag4	-.0006***	.0002	-.0006***	.0002	-.0006***	.0002
RoE_L1	.0005	.0038	.0007	.0039	.0133**	.0063
MBR_L1	.0000	.0000	.0000	.0000	.0000	.0000
Leverage_L1	.0000	.0000	.0000	.0000	-.0001*	.0001
Intangibles_L1	.0004	.0003	.0004	.0003	.0044**	.0021
Goodwill_L1	-.0004	.0003	-.0004	.0003	-.0066***	.0021
Firm Size_L1	-.0009*	.0005	-.0011	.0007	-.0115**	.0046
Cap. Ex._L1	.0001	.0002	.0002	.0002	.0018	.0012
Tobin's Q_L1	.0000	.0001	.0000	.0001	-.0001	.0003
CPI_L1	-.0002***	.0001	-.0002***	.0001	-.0001	.0001
Constant	.0595***	.0194	.0594***	.0194	.0682***	.0204
SubSector FE	No		Yes		No	
Firm FE	No		No		Yes	
Week FE	No		No		No	
N	5415		5415		5415	
Firms	95		95		95	
Lags	4		4		4	
F-Test	4.3916***		4.3313***		1.1642*	
R2-adjusted	.0142		.0069		.0231	

Table 15(2). Effect of investor attention by SGSV Combined (pre-Covid)

This table presents three regressions regarding the effect of investor attention during the pre-Covid period, measured by SGSV Combined data. Therefore, the sample period is January 2019 to February 2020. The regressions showed in this table are a subsector + week fixed effects, firm + week fixed effects, and subsector + firm + week fixed effects model. Moreover, the table includes 4 lags of the SGSV data, a lag of the dependent variable, and a lag of the control variables. The F-test indicates the significance of the entire model and the R2-adjusted is, compared to the R2, adjusted for the number of predictor variables in the model. ***, **, and *, indicate the 1%, 5%, and 10% significance levels, respectively.

Weekly Average	**SubSector + Week FE**		**Firm + Week FE**		**Firm + SubSector + Week FE**	
	Coefficient	**St. Error**	**Coefficient**	**St. Error**	**Coefficient**	**St. Error**
WA_L1	-.0259*	.014	-.0358**	.0141	-.0358**	.0141
Lag1	.0002	.0002	.0002	.0002	.0002	.0002
Lag2	.0001	.0002	.0001	.0002	.0001	.0002
Lag3	-.0002*	.0002	-.0002*	.0002	-.0002*	.0002
Lag4	-.0003*	.0002	-.0003*	.0002	-.0003*	.0002
RoE_L1	-.0037	.0033	-.0033	.0057	-.0033	.0057
MBR_L1	.0000	.0000	.0000	.0000	.0000	.0000
Leverage_L1	.0000	.0000	-.0001	.0001	-.0001	.0001
Intangibles_L1	.0003	.0003	.0021	.0017	.0021	.0017
Goodwill_L1	-.0003	.0003	-.0040**	.0018	-.004**	.0018
Firm Size_L1	-.0005	.0006	.006	.0045	.006	.0045
Cap. Ex._L1	.0001	.0002	.0016	.001	.0016	.001
Tobin's Q_L1	.0000	.0001	.0002	.0003	.0002	.0003
CPI_L1	-.0038***	.0000	-.0039***	.0002	-.0039***	.0002
Constant	.9618***	.0416	.9561***	.0466	.957***	0.465
SubSector FE	Yes		No		Yes	
Firm FE	No		Yes		Yes	
Week FE	Yes		Yes		Yes	
N	5415		5415		5415	
Firms	95		95		95	
Lags	4		4		4	
F-Test	35.8314		35.8694		16.4098	
R2-adjusted	.3066		.2992		.3039	

Based on the results in Table 15(1) and Table 15(2), it is decided which regressions will be performed in the remainder of this thesis. First of all, the regressions including the week fixed effects (Table 15(2)) show a sharp increase in the R2-adjusted and the significance of the entire model (F-test). Therefore, it was decided not to perform the regressions without week fixed effects (only subsector

fixed effects and only firm fixed effects) in the remainder of the study. However, the SGSV results do show significant (or less significant) results in the latter regressions, in contrast to the regressions that include week fixed effects. In contrast, Tables 15(1) and 15(2) only concern the period before the Covid-19 outbreak, so they include fewer observations. The regressions including the entire sample plus week fixed effects show that the variables regarding the SGSV data are significant. Therefore, it is concluded that the insignificant SGSV variables from Table 15(2) are not caused by the inclusion of week fixed effects, but possibly by a smaller sample. For this reason, the choice of regressions is based on the increase in R2-adjusted. In addition, regression 6 (subsector, firm, and week fixed effects) contains omitted variables due to the inclusion of both firm and subsector fixed effects. Therefore, it is decided to also exclude this regression in the next tables. Subsequently, the base regression (Pooled OLS), along with the two previously mentioned regressions, will also be shown for subsequent regressions. The tables in Appendix G, including the results of the SGSV by ticker and by company name measures, only contain the three above selected regressions.

Table 15, including the SGSV combined results, shows results which correspond to hypothesis 3a and 3b. The expectations include a positive effect of investor attention on stock returns in the short term (first 1-2 weeks) and a negative effect on the long term (after reversal point). The other measures for investor attention also show such a reversal point after 1 or 2 weeks (Appendix G). The coefficient of SGSV by Ticker becomes negative after 1 week. In addition, the coefficients of SGSV by Company Name becomes negative after 2 weeks. Therefore, a positive effect in the short term, a negative effect in the long term, and a reversal point are established.

5.5 Effect of investor attention during Covid-19

Hypothesis 4a and 4b focus on the period after the Covid-19 outbreak and a comparison to the results from hypothesis 3a and 3b. Therefore, the same regression analysis will be performed. Although, only months after the Covid-19 outbreak are included this time. This table and all subsequent tables only show the OLS, subsector plus week fixed effects, and firm plus week fixed effects regressions. Table 16 contains the results of the SGSV combined measure. The results of the SGSV by ticker and SGSV by company name can be found in Appendix H. The results of the regressions can be found below, in Table 16. First, the results of Table 16 will be discussed separately. Thereafter, the results of Table 15 and Table 16 will be compared to draw a conclusion for hypothesis 4a and 4b. The results from SGSV by Ticker and Company name act as a robustness check on the SGSV Combined data, again.

Table 16. Effect of investor attention by SGSV Combined (post-Covid)

This table presents three regressions regarding the effect of investor attention during the post-Covid period, measured by SGSV Combined data. Therefore, the sample period is February 2020 to April 2021. The regressions showed in this table are an OLS, subsector + week fixed effects, and firm + week fixed effects regression. Moreover, the table includes 4 lags of the SGSV data, a lag of the dependent variable, and a lag of the control variables. The F-test indicates the significance of the entire model and the R2-adjusted is, compared to the R2, adjusted for the number of predictor variables in the model. ***, **, and *, indicate the 1%, 5%, and 10% significance levels, respectively.

Weekly Average	**OLS**		**SubSector + Week FE**		**Firm + Week FE**	
	Coefficient	**St. Error**	**Coefficient**	**St. Error**	**Coefficient**	**St. Error**
WA_L1	-.1180***	.0127	-.1083***	.013	-.1201***	.0131
Lag1	.0006**	.0002	.0013***	.0002	.0013***	.0002
Lag2	-.0012***	.0003	-.0011***	.0002	-.0011***	.0002
Lag3	-.0004	.0003	-.0006***	.0002	-.0007***	.0002
Lag4	-.0004	.0002	-.0003*	.0002	-.0004*	.0002
RoE_L1	-.0003	.003	-.0006	.0028	.0037	.0051
MBR_L1	.0000	.0000	.0000	.0000	0	0
Leverage_L1	.0000	.0001	.0000	.0000	0	.0001
Intangibles_L1	-.0007*	.0004	-.0005	.0003	-.002	.0021
Goodwill_L1	.0005	.0004	.0004	.0003	.0037	.0023
Firm Size_L1	.0008	.0007	.0010	.0006	-.0017	.0044
Cap. Ex._L1	-.0002	.0002	-.0004	.0002	-.0014*	.0008
Tobin's Q_L1	.0001	.0001	.0001*	.0001	.0003	.0003
CPI_L1	-.0002***	.0001	-.0008***	.0002	-.0008***	.0002
Constant	.0554***	.0157	.2161***	.0535	.2150***	.0541
SubSector FE	No		Yes		No	
Firm FE	No		No		Yes	
Week FE	No		Yes		Yes	
N	5795		5795		5795	
Firms	95		95		95	
Lags	4		4		4	
F-Test	10.7822***		50.9811***		50.9772***	
R2-adjusted	.0231		.3858		.3802	

The coefficients of the lags in Table 16 show a similar pattern as the results in Table 15, the pre-Covid period. Although, during the Covid-19 period and in contrast with the pre-Covid period, the reversal point occurs after 1 week for the SGSV combined measurement (instead of 2). However, the results are still in line with the expectations and the hypothesis on the short and long term. The lag (week) of

the reversal point is significant for each measurement again. Although, the other SGSV lags are less or not significant.

By comparing Table 15 and Table 16, a conclusion can be drawn on hypothesis 4a and 4b. The expectations were that the effect of investor attention on stock returns would be more positive in both the short (4a) and long (4b) term. In other words, the coefficients of the lags of SGSV measurements in Table 16 should be higher/more positive than in Table 15. Although the coefficients turn negative after the reversal point, the expectation is a more positive coefficient during Covid-19. First, the results of the OLS regression will be compared. The results show more positive effects in week 1 (.0006 > .0003), week 3 (-.0004 > -.0005), and week 4 (-.0004 > -.0006). However, the regressions including the fixed effects only show a more positive results for week 1 (.0013 > .0002). The other lags (weeks) are more positive during the period before the Covid-19 outbreak. Therefore, based on the SGSV combined measure, it can be concluded that the effect of investor attention is more positive during Covid-19 in the short term, before the reversal point. This corresponds to the expectation of hypothesis 4a. In contrast, the effect of investor attention in the long term seems to be more positive in the period before Covid-19, which does not match the expectation of hypothesis 4b. Appendix G and Appendix H show the result of the SGSV by ticker and by company name measures. Table 35(1) and Table 36(1) compare the results of SGSV by Ticker data and Table 35(2) and Table 36(2) compare the results of SGSV by company name data. The results are similar to the results described above for SGSV combined data. Week 1 is more positive during Covid-19, while the other weeks (longer term) show a more positive effect of investor attention in the long term.

5.6 Amount of investor attention (directly vs indirectly related firms)

In this section, a similar analysis will be performed as for hypothesis 2. However, a distinction between directly related and not directly related firms to Covid-19 is made for hypothesis 5a. The hypothesis states the following: '*During Covid-19 (from March 2020), healthcare companies directly related to Covid-19 receive more attention than healthcare companies not directly related to Covid-19.'* Therefore, the descriptive statistics of both groups of companies will be compared. Appendix D contains a list with all directly related companies. In Table 17 the amount of investor attention for both groups can be found. The investor attention is given for all measures of SGSV, similar to the comparison in hypothesis 1.

Table 17. SGSVs Directly related vs not directly related

The amount of investor attention, based on the SGSV Combined measure, is shown in Table 17. In addition, the results are divided into 4 rows based on relatedness of the firm to Covid-19 and pre-Covid/during Covid-19.

	SGSV by Ticker	SGSV by Company Name	SGSV Combined
Pre-Covid directly related	-.277	-.115	-.196
Pre-Covid not directly related	.120	.197	.159
During Covid directly related	.283	.066	.175
During Covid not directly related	-.069	-.149	-.109

Table 17 shows that companies directly related to the Covid-19 crisis/virus received much more attention during the period after the Covid outbreak compared to the pre-Covid period. Since standardized GSV data is used, the search data per company is valued in relation to the search data to the same company on the other days during the sample period. Therefore, the negative values in the first row have to be compared to the values in the third row (same companies). In addition, the negative value does not mean that there was no attention, but this is a relative value with respect to the value from row 3. Table 17 clearly shows that directly related companies received more attention during Covid-19, while that actually decreased for the unrelated companies (row 4 relative to row 2).

5.7 Effect of investor attention (directly vs indirectly related firms)

Section 5.7 will examine the difference in terms of effect of investor attention between directly and not directly related firms. A pooled OLS and two fixed effects regression will be performed again. The results of these regressions can be found in Table 18. This table only shows the combined SGSV data for both the directly related and not directly related companies. The results of the SGSV by ticker and company measures can be found in Appendix I.

Table 18(1). Effect of directly related companies (SGSV Combined)

This table presents three regressions regarding the effect of investor attention measured by SGSV Combined data. It only includes firms that are considered as directly related to the Covid-19 pandemic. Therefore, the sample period is January 2019 to April 2021. The regressions showed in this table are an OLS, subsector + week fixed effects, and firm + week fixed effects regression. Moreover, the table includes 4 lags of the SGSV data, a lag of the dependent variable, and a lag of the control variables. The F-test indicates the significance of the entire model and the R2-adjusted is, compared to the R2, adjusted for the number of predictor variables in the model. ***, **, and *, indicate the 1%, 5%, and 10% significance levels, respectively.

Weekly Average	**OLS**		**SubSector + Week FE**		**Firm + Week FE**	
	Coefficient	**St. Error**	**Coefficient**	**St. Error**	**Coefficient**	**St. Error**
WA_L1	-.0520***	.0156	-.0466***	.0158	-.0572***	.0158
Lag1	.0000	.0002	.0004*	.0002	.0004*	.0002
Lag2	-.0004*	.0003	-.0003***	.0002	-.0003***	.0002
Lag3	-.0003	.0003	-.0002**	.0002	-.0002**	.0002
Lag4	-.0000	.0002	-.0000	.0002	-.0000	.0002
RoE_L1	-.0005	.0030	-.0008	.0027	.0000	.004
MBR_L1	.0000	.0000	.0000	.0000	.0000	.0000
Leverage_L1	.0000	.0000	.0000	.0000	-.0001	.0001
Intangibles_L1	-.0004	.0009	-.001	.0009	-.0033	.0031
Goodwill_L1	.0003	.0009	.001	.0009	.0028	.0032
Firm Size_L1	.0013	.0007	.0016**	.0007	.0086***	.0025
Cap. Ex._L1	-.0006**	.0003	-.0009***	.0003	-.0017**	.0008
Tobin's Q_L1	.0001	.0001	.0001	.0001	.0004**	.0002
CPI_L1	.0000	.0001	-.0006***	.0001	-.0007***	.0001
Constant	.0069	.0138	.1540***	.0377	.1513***	.0379
SubSector FE	No		Yes		No	
Firm FE	No		No		Yes	
Week FE	No		Yes		Yes	
N	4130		4130		4130	
Firms	35		35		35	
Lags	4		4		4	

F-Test	2.4190***	16.3777***	16.5177***
R2-adjusted	.0048	.3254	.3245

Table 18(2). Effect of companies not directly related (SGSV Combined)

This table presents three regressions regarding the effect of investor attention measured by SGSV Combined data. It only includes firms that are considered as not directly related to the Covid-19 pandemic. Therefore, the sample period is January 2019 to April 2021. The regressions showed in this table are an OLS, subsector + week fixed effects, and firm + week fixed effects regression. Moreover, the table includes 4 lags of the SGSV data, a lag of the dependent variable, and a lag of the control variables. The F-test indicates the significance of the entire model and the R2-adjusted is, compared to the R2, adjusted for the number of predictor variables in the model. ***, **, and *, indicate the 1%, 5%, and 10% significance levels, respectively.

Weekly Average	**OLS**		**SubSector + Week FE**		**Firm + Week FE**	
	Coefficient	**St. Error**	**Coefficient**	**St. Error**	**Coefficient**	**St. Error**
WA_L1	-.1198***	.0118	-.0889***	.0119	-.0933***	.012
Lag1	.0006***	.0002	.0010***	.0002	.0010***	.0002
Lag2	-.0007***	.0002	-.0005**	.0002	-.0005**	.0002
Lag3	-.0005**	.0002	-.0003*	.0002	-.0004*	.0002
Lag4	-.0006***	.0002	-.0004**	.0002	-.0005**	.0002
RoE_L1	-.0023	.0033	-.0005	.0031	.0004	.0038
MBR_L1	.0000	.0000	.0000	.0000	.0000	.0000
Leverage_L1	.0000	.0000	.0000	.0000	.0000	.0000
Intangibles_L1	.0001	.0003	.0003	.0003	.0005	.0009
Goodwill_L1	.0000	.0003	-.0002	.0002	.0004	.0009
Firm Size_L1	-.0005	.0006	-.0004	.0006	-.0034	.0032
Cap. Ex._L1	.0000	.0002	-.0001	.0002	.0001	.0005
Tobin's Q_L1	.0001*	.0001	.0002***	.0001	.0004**	.0002
CPI_L1	.0000	.0000	-.0004***	.0001	-.0004***	.0001
Constant	-.0009	.0123	.0947***	.0325	.0982***	.0329
SubSector FE	No		Yes		No	
Firm FE	No		No		Yes	
Week FE	No		Yes		Yes	
N	7080		7080		7080	
Firms	60		60		60	
Lags	4		4		4	
F-Test	11.2155***		36.6253		36.5422***	
R2-adjusted	.0198		.3950		.3919	

Hypothesis 5b can be answered by the results of Tables 18(1) and 18(2). It appears that the coefficients of the directly related firms are more positive in the long term, or at least after the reversal

point. Again, the reversal point is already after 1 week, for both directly related and not directly related firms. In this first week, the effect of investor attention is more positive for companies not directly related (.0006/.0010 > .0000/.0004). As a result, no unequivocal conclusion can be drawn about the short term (1-2 weeks), as week 2 shows opposite results. The coefficients of week 2 to 4 are more positive for related companies. Therefore, hypothesis 5b can be partly confirmed because investor attention has a more positive effect in the longer term. Although, the results of lag 4 of the related firms are less or not significant which may be caused by the lower number of companies and observations. This makes it difficult to draw a definite conclusion, even on the long term. The results of the other measures (Appendix I) show similar results as the SGSV combined measure.

5.8 Effect of investor attention for directly related firms (pre-covid vs during covid)

Subsequently, a comparison is made between the effect of investor attention for directly related companies in the pre-Covid period and after the Covid outbreak period. Tables 19(1) and 19(2) show the results of both regressions, in which only the combined SGSV data is included again. To conclude on hypothesis 5c, only the results of the SGSV combined measure are analyzed because the other two measures showed almost identical results in the previous regressions.

Table 19(1). Effect investor attention directly related firms pre-Covid (SGSV Combined)

This table presents three regressions regarding the effect of investor attention measured by SGSV Combined data during the pre-Covid period. It only includes firms that are considered as directly related to the Covid-19 pandemic. Therefore, the sample period is January 2019 to February 2020. The regressions showed in this table are an OLS, subsector + week fixed effects, and firm + week fixed effects regression. Moreover, the table includes 4 lags of the SGSV data, a lag of the dependent variable, and a lag of the control variables. The F-test indicates the significance of the entire model and the R2-adjusted is, compared to the R2, adjusted for the number of predictor variables in the model. ***, **, and *, indicate the 1%, 5%, and 10% significance levels, respectively.

Weekly Average	**OLS**		**SubSector + Week FE**		**Firm + Week FE**	
	Coefficient	**St. Error**	**Coefficient**	**St. Error**	**Coefficient**	**St. Error**
WA_L1	-.0919***	.0242	-.0330*	.0242	-.0445*	.0243
Lag1	-.0003	.0003	-.0006*	.0003	-.0006*	.0003
Lag2	.0002	.0003	.0001	.0003	.0001	.0003
Lag3	-.0008**	.0003	-.0004*	.0003	-.0004*	.0003
Lag4	-.0004	.0003	-.0001	.0003	-.0001	.0003
RoE_L1	-.0026	.0049	-.0047	.0044	-.0164**	.0073
MBR_L1	.0000	.0000	.0000	.0000	.0000	.0000
Leverage_L1	.0000	.0001	-.0001	.0001	.0000	.0001
Intangibles_L1	.0008	.0011	-.0003	.0011	-.0031	.0079
Goodwill_L1	-.0008	.0011	.0004	.0012	-.0009	.0072
Firm Size_L1	-.0003	.001	-.0001	.0011	.0213***	.0056
Cap. Ex._L1	-.0001	.0004	-.0001	.0005	-.0002	.0020
Tobin's Q_L1	-.0002	.0001	-.0001	.0001	-.0001	.0005
CPI_L1	-.0003***	.0001	-.0033***	.0003	-.0035***	.0003
Constant	.0926***	.0284	.8307***	.0682	.8379***	.0739
SubSector FE	No		Yes		No	
Firm FE	No		No		Yes	
Week FE	No		Yes		Yes	
N	1995		1995		1995	
Firms	35		35		35	
Lags	4		4		4	
F-Test	3.0096***		13.6820***		14.0410***	
R2-adjusted	.0139		.3033		.3028	

Table 19(2). Effect investor attention directly related firms post-Covid (SGSV Combined)

This table presents three regressions regarding the effect of investor attention measured by SGSV Combined data during the Covid-19 period. It only includes firms that are considered as directly related to the Covid-19 pandemic. Therefore, the sample period is February 2020 to April 2021. The regressions showed in this table are an OLS, subsector + week fixed effects, and firm + week fixed effects regression. Moreover, the table includes 4 lags of the SGSV data, a lag of the dependent variable, and a lag of the control variables. The F-test indicates the significance of the entire model and the R2-adjusted is, compared to the R2, adjusted for the number of predictor variables in the model. ***, **, and *, indicate the 1%, 5%, and 10% significance levels, respectively.

Weekly Average	**OLS**		**SubSector + Week FE**		**Firm + Week FE**	
	Coefficient	**St. Error**	**Coefficient**	**St. Error**	**Coefficient**	**St. Error**
WA_L1	-.0516**	.0210	-.0577***	.0215	-.0696***	.0216
Lag1	.0002	.0003	.001***	.0003	.0009***	.0003
Lag2	-.0010**	.0004	-.0011***	.0003	-.0012***	.0003
Lag3	-.0001	.0004	-.0004	.0003	-.0004	.0003
Lag4	-.0001	.0003	-.0000	.0003	-.0001	.0003
RoE_L1	.0016	.0041	-.0043	.0046	.0110	.0103
MBR_L1	.0000	.0000	.0001	.0000	.0000	.0001
Leverage_L1	.0000	.0001	-.0001	.0001	-.0001	.0002
Intangibles_L1	-.0003	.0014	-.0021	.0016	-.0021	.0116
Goodwill_L1	.0001	.0015	.0019	.0017	.0026	.0127
Firm Size_L1	.0015	.0011	.0025**	.0011	.0055	.0073
Cap. Ex._L1	-.0007	.0005	-.0015***	.0005	-.0037**	.0016
Tobin's Q_L1	.0001	.0001	-.0001	.0002	.0006	.0005
CPI_L1	-.0003***	.0001	-.0014***	.0003	-.0014***	.0003
Constant	.0874***	.0232	.3733***	.0812	.3692***	.0823
SubSector FE	No		Yes		No	
Firm FE	No		No		Yes	
Week FE	No		Yes		Yes	
N	2135		2135		2135	
Firms	35		35		35	
Lags	4		4		4	
F-Test	3.1206***		15.6869***		15.6619***	
R2-adjusted	.0137		.3329		.3269	

The results of Tables 19(1) and 19(2) will be discussed below. First, it is remarkable that the first week of the pre-Covid period shows a negative coefficient. Although, this could be explained by the fact that the sample of pre-Covid and directly related companies is limited in terms of observations. This could potentially influence the results. In addition, the first lag results are not significant. However, after 2 weeks (pre-Covid) and 1 week (post-Covid outbreak) there is a reversal to negative

coefficients. Therefore, the timing of the reversal corresponds to the expectations. Both reversal points (pre- and post-Covid) are significant. In addition, the effect on the long term (after reversal point) shows remarkable results. The outcomes of lag 3 and 4 for both the pre- and post-Covid outbreak period are almost identical. Unfortunately, both results in the short and longer term are not completely unambiguous. Therefore, a general conclusion cannot be drawn about both terms. Earlier, the low number of observations was already mentioned regarding the related companies. In addition, this number has been further reduced because a distinction is made between the two periods. This may affect the significance of the SGSV data. Only the reversal points in both tables (lag 2 and lag 3) show significance for all regressions.

5.9 Amount of investor attention (Firm size and relatedness)

Hypothesis 6a examine the amount of investor attention for different combinations of firm size and relatedness to Covid-19. In Table 20, the amount of investor attention is studied on the basis of a comparison of the descriptive statistics of these different combinations. Only the period after the Covid-19 outbreak will be included to perform this analysis. Based on the existing literature, the expectation is that larger firms (higher firm size) receive more attention. Larger companies are more often in the news and are more known to the public and investors. In addition, due to the health nature of the crisis, it is expected that companies that are directly related to Covid-19 will receive more attention. The results can be found in Table 20. Table 21 contains the regression results to measure the effect of the attention for these different combinations of firm size and relatedness to Covid-19. Quantile 4 contains the firms with the largest firm size.

Table 20. Investor attention based on Firm Size and Relatedness (during Covid-19)

This table shows the SGSV results for all three measures. In addition, the results are divided into 8 groups based on 4 quantiles of firm size and 2 option regarding to relatedness to Covid-19 (related/unrelated).

	SGSV by Ticker	SGSV by Company Name	SGSV Combined
Q4 + related	.381	.082	.231
Q4 + unrelated	.092	-.145	-.026
Q3 + related	.275	.131	.203
Q3 + unrelated	-.058	-.178	-.118
Q2 + related	.239	-.024	.108
Q2 + unrelated	-.148	-.218	-.182
Q1 + related	.121	.015	.068
Q1 + unrelated	-.100	-.053	-.076

Table 20 matches the expectations. First, the highest amount of investor attention is received by related firms in the highest firm size quantile (Q4), based on SGSV by Ticker (.381) and SGSV Combined (.231). Although, based on the results by company name, related firms in Q3 received the most attention. However, Q3 includes firms with the second highest firm size which can explain the deviation. Second, the SGSV Combined results show a gradual decrease in amount of investor attention per quantile (both for related as unrelated firms). The only deviation is the fact that firms in Q2 received less attention than firms in Q1 (for all measures). In general, it can be concluded that the results of Table 20 correspond to hypothesis 6a.

Table 21(1), 21(2), 21(3) and 21(4) can be found below. It includes 4 combinations, as Q1 + Q2 and Q3 + Q4 are merged together. This leaves two options in terms of firm size and two options in terms of relatedness. Therefore, four separate regressions have been performed. The results are based on the SGSV Combined. In total, 4 lags will be analyzed. Again, only the SGSV combined measure is analyzed for this specific topic.

Table 21(1). Effect investor attention firm size + relatedness: Q3+Q4 related (during Covid)

This table presents three regressions regarding the effect of investor attention measured by SGSV Combined data during the Covid-19 period. It only includes firms that are considered as directly related to the Covid-19 pandemic and belongs to the largest firm sizes (Q3+Q4). Therefore, the sample period is February 2020 to April 2021. The regressions showed in this table are an OLS, subsector + week fixed effects, and firm + week fixed effects regression. Moreover, the table includes 4 lags of the SGSV data, a lag of the dependent variable, and a lag of the control variables. The F-test indicates the significance of the entire model and the R2-adjusted is, compared to the R2, adjusted for the number of predictor variables in the model. ***, **, and *, indicate the 1%, 5%, and 10% significance levels, respectively.

Weekly Average	**OLS**		**SubSector + Week FE**		**Firm + Week FE**	
	Coefficient	**St. Error**	**Coefficient**	**St. Error**	**Coefficient**	**St. Error**
WA_L1	-.0822***	.0249	-.1035***	.0259	-.1127***	.0259
Lag1	.0002	.0003	.0009***	.0003	.0009***	.0003
Lag2	-.0004	.0004	-.0006**	.0003	-.0007**	.0003
Lag3	-.0001	.0004	-.0002*	.0003	-.0003*	.0003
Lag4	-.0001	.0003	-.0001	.0003	-.0003	.0003
RoE_L1	.0024	.0038	-.0009	.0045	.0060	.0086
MBR_L1	.0000	.0000	.0000	.0000	.000	.0001
Leverage_L1	-.0001	.0001	-.0001	.0001	-.0001	.0001
Intangibles_L1	-.0009	.0015	-.0015	.0014	-.003	.0100
Goodwill_L1	.0009	.0016	.0014	.0014	.0044	.0108
Firm Size_L1	.0004	.0013	.0012	.0013	-.0119	.0086
Cap. Ex._L1	-.0003	.0005	-.0007	.0006	-.0021	.0022
Tobin's Q_L1	.0001	.0002	.0000	.0002	-.0003	.0008
CPI_L1	-.0002**	.0001	-.0012***	.0003	-.0011***	.0003
Constant	.0604**	.0234	.3271***	.0761	.3518***	.0779
SubSector FE	No		Yes		No	
Firm FE	No		No		Yes	
Week FE	No		Yes		Yes	
N	1498		1498		1498	
Firms	24		24		24	
Lags	4		4		4	
F-Test	1.3886		17.2675***		17.4659***	
R2-adjusted	.0036		.4411		.4402	

Table 21(2). Effect investor attention firm size + relatedness: Q3+Q4 unrelated (during Covid)

This table presents three regressions regarding the effect of investor attention measured by SGSV Combined data during the Covid-19 period. It only includes firms that are considered as not directly related to the Covid-19 pandemic and belongs to the largest firm sizes (Q3+Q4). Therefore, the sample period is February 2020 to April 2021. The regressions showed in this table are an OLS, subsector + week fixed effects, and firm + week fixed effects regression. Moreover, the table includes 4 lags of the SGSV data, a lag of the dependent variable, and a lag of the control variables. The F-test indicates the significance of the entire model and the R2-adjusted is, compared to the R2, adjusted for the number of predictor variables in the model. ***, **, and *, indicate the 1%, 5%, and 10% significance levels, respectively.

Weekly Average	**OLS**		**SubSector + Week FE**		**Firm + Week FE**	
	Coefficient	**St. Error**	**Coefficient**	**St. Error**	**Coefficient**	**St. Error**
WA_L1	-.1993***	.0238	-.2272***	.0249	-.2297***	.0250
Lag1	.0001	.0005	.0007*	.0004	.0006	.0004
Lag2	-.0011**	.0005	-.0004	.0004	-.0004	.0004
Lag3	-.0006	.0005	-.0002	.0004	-.0003	.0004
Lag4	.0006	.0005	-.0001	.0004	-.0001	.0004
RoE_L1	.0033	.0110	.0082	.0095	.0028	.0191
MBR_L1	.0000	.0000	.0000	.0000	.0000	.0000
Leverage_L1	-.0001	.0001	-.0001	.0001	-.0001	.0001
Intangibles_L1	-.0011	.0012	.0000	.0014	.0046	.0133
Goodwill_L1	.001	.0011	.0000	.0012	-.0057	.0147
Firm Size_L1	.0013	.0024	.0016	.0034	-.0072	.0183
Cap. Ex._L1	-.0004	.0006	-.0004	.0007	.0013	.0036
Tobin's Q_L1	-.0001	.0003	.0001	.0003	.0004	.0015
CPI_L1	.0000	.0001	-.0002	.0003	-.0002	.0004
Constant	-.0053	.0311	.0603	.0907	.0884	.1066
SubSector FE	No		Yes		No	
Firm FE	No		No		Yes	
Week FE	No		Yes		Yes	
N	1593		1593		1593	
Firms	29		29		29	
Lags	4		4		4	
F-Test	5.8714***		28.9451***		28.6526***	
R2-adjusted	.0411		.5606		.5556	

Table 21(3). Effect investor attention firm size + relatedness: Q1+Q2 related (during Covid)

This table presents three regressions regarding the effect of investor attention measured by SGSV Combined data during the Covid-19 period. It only includes firms that are considered as directly related to the Covid-19 pandemic and belongs to the largest firm sizes (Q1+Q2). Therefore, the sample period is February 2020 to April 2021. The regressions showed in this table are an OLS, subsector + week fixed effects, and firm + week fixed effects regression. Moreover, the table includes 4 lags of the SGSV data, a lag of the dependent variable, and a lag of the control variables. The F-test indicates the significance of the entire model and the R2-adjusted is, compared to the R2, adjusted for the number of predictor variables in the model. ***, **, and *, indicate the 1%, 5%, and 10% significance levels, respectively.

Weekly Average	**OLS**		**SubSector + Week FE**		**Firm + Week FE**	
	Coefficient	**St. Error**	**Coefficient**	**St. Error**	**Coefficient**	**St. Error**
WA_L1	-.0405	.039	-.0516	.0408	-.0588	.0410
Lag1	.0000	.0009	.0007	.0008	.0007	.0009
Lag2	-.0027***	.0009	-.0027***	.0009	-.0028***	.0009
Lag3	-.0007	.0009	-.0013	.0009	-.0015	.0009
Lag4	-.0002	.0009	-.0005	.0009	-.0003	.0009
RoE_L1	.044	.0338	.0255	.0365	.0641	.6804
MBR_L1	-.0001	.0002	-.0001	.0002	-.0016	.0018
Leverage_L1	.0009	.0022	.0000	.0030	.0048	.0092
Intangibles_L1	.0051	.0079	-.0072	.0094	-.0113	.0667
Goodwill_L1	-.0059	.0085	.008	.0102	.0376	.0835
Firm Size_L1	.009**	.0045	.0211***	.0066	-.0761	.2012
Cap. Ex._L1	-.0028*	.0016	-.007***	.0025	-.0039	.0105
Tobin's Q_L1	.0006	.0008	.0010	.0009	.0034	.0030
CPI_L1	-.0005**	.0002	-.0015*	.0008	-.0011	.0018
Constant	.1142*	.0594	.3347	.2096	.4344	.3055
SubSector FE	No		Yes		No	
Firm FE	No		No		Yes	
Week FE	No		Yes		Yes	
N	637		637		637	
Firms	11		11		11	
Lags	4		4		4	
F-Test	2.2926***		3.8084***		3.8219***	
R2-adjusted	.0277		.2402		.2356	

Table 21(4). Effect investor attention firm size + relatedness: Q1+Q2 unrelated (during Covid)

This table presents three regressions regarding the effect of investor attention measured by SGSV Combined data during the Covid-19 period. It only includes firms that are considered as not directly related to the Covid-19 pandemic and belongs to the largest firm sizes (Q1+Q2). Therefore, the sample period is February 2020 to April 2021. The regressions showed in this table are an OLS, subsector + week fixed effects, and firm + week fixed effects regression. Moreover, the table includes 4 lags of the SGSV data, a lag of the dependent variable, and a lag of the control variables. The F-test indicates the significance of the entire model and the R2-adjusted is, compared to the R2, adjusted for the number of predictor variables in the model. ***, **, and *, indicate the 1%, 5%, and 10% significance levels, respectively.

Weekly Average	**OLS**		**SubSector + Week FE**		**Firm + Week FE**	
	Coefficient	**St. Error**	**Coefficient**	**St. Error**	**Coefficient**	**St. Error**
WA_L1	-.1159***	.0214	-.088***	.0223	-.0989***	.0224
Lag1	.0012***	.0005	.0019***	.0004	.0019***	.0004
Lag2	-.0017***	.0005	-.0011***	.0004	-.0010**	.0004
Lag3	-.0009*	.0005	-.0010**	.0004	-.0010**	.0004
Lag4	-.0017***	.0005	-.0008**	.0004	-.0009**	.0004
RoE_L1	-.0066	.0053	-.0037	.0061	-.0008	.0087
MBR_L1	.0000	.0000	.0000	.0000	.0000	.0001
Leverage_L1	-.0003	.0004	-.0005	.0005	-.0005	.0007
Intangibles_L1	-.0006	.0006	-.0003	.0005	-.0021	.0028
Goodwill_L1	.0005	.0005	.0003	.0005	.0047	.0029
Firm Size_L1	-.0001	.0018	-.0004	.0017	-.0022	.0094
Cap. Ex._L1	-.0001	.0004	-.0003	.0005	-.0005	.0013
Tobin's Q_L1	.0002	.0001	.0002	.0001	.0002	.0005
CPI_L1	-.0003**	.0001	-.0005	.0004	-.0005	.0004
Constant	.0683**	.0294	.1295	.0993	.1338	.1016
SubSector FE	No		Yes		No	
Firm FE	No		No		Yes	
Week FE	No		Yes		Yes	
N	2067		2067		2067	
Firms	35		35		35	
Lags	4		4		4	
F-Test	5.6660***		18.1554***		18.1664***	
R2-adjusted	.0306		.3759		.3711	

The results of Tables 21(1), 21(2), 21(3), and 21(4) show similarities with expectations in some respects. It was expected that the results for the largest and related companies (Table 21(1)) would show that the effect of investor attention is greatest for this group. However, the results from week 1 are contradictory, as the smallest and unrelated companies (Table 21(4)) have by far the highest effect here. On the other hand, the reversal point is again after week 1 and weeks 2 to 4 clearly show that both groups of the largest companies (both related and unrelated) have a much more positive effect

than the companies in Tables 21(3) and 21(4). Therefore, it can be concluded that firm size has a significant positive effect on the effect of investor attention in the longer term, or after the reversal point (lag 2 to 4). It is more difficult to draw a firm conclusion about the effect of relatedness from the four Tables 21. Table 21(4) (unrelated) does not show an unequivocally lower effect compared to Table 21(3) (related). The same applies to Table 21(2) compared to 21(1). However, this has been demonstrated in Tables 18(1) and 18(2). There could be several reasons for the results in Tables 21. First, the selection of firms based on relatedness is not proportional, while firm size is. For example, Table 21(4) has many more observations than Table 21(3), which may affect the results and significance.

5.10 Investor attention per subsector (SIC code)

Based on the results of a pooled OLS regression, the individual effect per subsector can be examined. There is made a distinction between 9 subsectors: Pharmaceuticals, Managed Healthcare, Life Sciences Tools & Services, Healthcare Supplies, Healthcare Services, Healthcare Facilities, Healthcare Equipment, Healthcare Distribution and Biotechnology. Since all dummies are included in the regression, there is no constant term that carries a particular effect. Again, the period before and after the Covid-19 outbreak will be compared to each other.

Table 22. Effect investor attention per subsector (pre- vs post-Covid outbreak)

This table presents two pooled OLS regressions regarding the effect of investor attention measured by SGSV Combined data before and during the Covid-19 period. Therefore, the sample period for the first regression is January 2019 to February 2020 and the sample period for the second regression is February 2020 to April 2021. Moreover, the table includes 4 lags of the SGSV data, a lag of the dependent variable, and a lag of the control variables. In addition, the 9 subsectors are shown separately. The F-test indicates the significance of the entire model and the R2-adjusted is, compared to the R2, adjusted for the number of predictor variables in the model. ***, **, and *, indicate the 1%, 5%, and 10% significance levels, respectively.

Weekly Average	**OLS (pre-Covid)**		**OLS (during Covid)**	
	Coefficient	**St. Error**	**Coefficient**	**St. Error**
WA_L1	-.0711***	.0144	-.1183	.0127
Lag1	.0003*	.0002	.0006	.0002
Lag2	.0001	.0002	-.0012	.0003
Lag3	-.0005**	.0002	-.0004	.0003
Lag4	-.0006***	.0002	-.0004	.0002
RoE_L1	.0007	.0039	-.001	.0036
MBR_L1	.0000	.0000	.0000	.0000
Leverage_L1	.0000	.0000	.0000	.0001
Intangibles_L1	.0004	.0003	-.0006	.0004
Goodwill_L1	-.0004	.0003	.0004	.0004
Firm Size_L1	-.0011	.0007	.0011	.0008
Cap. Ex._L1	.0002	.0002	-.0003	.0003

Tobin's Q_L1	.0000	.0001	.0001	.0001
CPI_L1	-.0002***	.0001	-.0002	.0001
Pharmaceuticals	.0003	.0011	-.0003	.0014
Managed HC	.0004	.0011	-.0007	.0014
Life Sciences T&S	.0008	.0010	-.0001	.0014
HC Supplies	.0000	.0011	-.0003	.0014
HC Services	.0006	.0011	-.0002	.0014
HC Facilities	.0002	.0012	.0007	.0015
HC Equipment	.0004	.0010	-.0006	.0013
HC Distributors	.0003	.0011	-.0006	.0015
Biotechnology	.0005	.0010	-.0003	.0013
SubSector FE	Yes		Yes	
Firm FE	No		No	
Week FE	No		No	
N	5415		5795	
Firms	95		95	
Lags	4		4	
F-Test	2.7610***		6.6621***	
R2-adjusted	.0074		.0220	

Table 22 shows the individual coefficients of the nine different subsectors for both periods before and after the Covid-19 outbreak measured by a pooled OLS regression. Immediately striking is the fact that the coefficients in the pre-Covid period are all positive and changes to negative values in the period after the outbreak. Therefore, it can be concluded that companies from all sectors experienced a positive effect of investor attention on stock returns during the pre-Covid period. However, after this outbreak of Covid-19, this has completely turned into a negative effect. An analysis of the individual subsectors shows that healthcare facilities group companies perform best, given the only positive value during the period after the Covid-19 outbreak (.0006).

6. Discussion

The aim of this thesis is to establish the relationship between investor attention by Google Trends data and stock performance of US healthcare companies during Covid-19. To do this, a comparison is made between the period before and after the Covid-19 outbreak. The discussion focuses on the limitations of this research, on which recommendations for future research are based. Subsequently, in chapter 7 an overall conclusion will be drawn on the basis of the results.

First of all, this thesis focuses only on US companies. As a result, a few companies that have played a major role in the fight against the Coronavirus are excluded from the study. An example is AstraZeneca, the developer of one of the few vaccines that have been widely used. The main reason for the choice of US companies is that Google Trends data can only be downloaded per country or worldwide. This allows search results from all over the US to be included. However, European searches are only possible per country. If companies all over the world were included in this research, global search data would be required. However, global search data increases the chance of searches for tickers or company names that have different meanings in other parts of the world. Therefore, the choice was made to focus on the US, although this poses a limitation of this study. Future research can conduct a similar study that takes multiple parts of the world into account. A solution to the limitation described above can be another source of search data.

Subsequently, this research focuses on weekly data. This is common in the existing literature that examines the relationship between investor attention and stock performance. In addition, daily data is more difficult to collect via Google Trends, since it can only be downloaded for shorter periods at once. Due to time limitations, it was decided to not include the daily data as an addition (robustness check) on the current study. However, daily data can provide more insight into the price pressure effect than weekly data. Weekly data sufficed for this study in the search for the effect of investor attention on stock returns and the associated reversal point. This reversal point was established in the existing literature after 1 to 2 weeks. Future research can extend this research by using daily data.

A third limitation of this thesis is the fact that only the public healthcare sector is included. The US healthcare sector is very diverse, in which the private sector also has a large share. The existing literature showed how closely the two sectors are linked, such as the many public-private partnerships. Unfortunately, due to the availability of data and the aim of this study, only the public sector could be investigated. The private sector generally contains smaller and more innovative companies (growth companies). This may provide a better comparison with the large established companies (S&P 500) than the mid-cap companies from the S&P 400. A recommendation for follow-up research is involving the private sector, if relevant data is available.

Thereafter, the number of companies in the sample that is directly related to the coronavirus is smaller than the number of companies not directly related to Covid-19 (35 vs 60). As a result, the number of observations of both regressions differ. Therefore, the comparison between the two groups based on a regression analysis may be slightly affected in reliability and validity. This ratio is distorted

because the company selection is based on certain indices (S&P 500 and 400 mid-cap). Future research could adjust the selection of companies accordingly to the right proportion and not make it dependent on an already existing selection of companies.

Finally, a larger sample size could have increased the number of observations, which may have a positive influence on the reliability. However, the period before and after the Covid-19 outbreak had to be equal in length in order to perform equal regressions. In addition, the availability of data needed to be dealt with. This resulted in the current sample selection. Future research may increase both periods (before and after Covid-19 outbreak) if more data from the period during Covid-19 becomes available.

7. Conclusion

Finally, this section will conclude on the results of this study as described in chapter 5. At the end of the conclusion**,** in Table 23, it is indicated for each hypothesis whether it is accepted or rejected. The aim of this thesis was to examine the relation between investor attention, measured by Google Trends data, and stock performance of US healthcare companies during Covid-19. In addition, the focus was on the difference between a period before Covid-19 (Jan 2019 to Feb 2020) and a period during Covid-19 (Mar 2020 to May 2021), as stated in the research question:

'What is the effect of GSV on healthcare stocks returns during Covid-19 compared to the pre-Covid period?'

First of all, the direction of causality was established by a Granger Causality test. This showed that investor attention Granger-cause stock returns at a 1% level for the combined SGSV data. In addition, robustness checks were performed with SGSV by ticker and SGSV by company name data, from which the same results emerged.

Secondly, an analysis was performed of the descriptive statistics of the SGSV data for both the S&P 500 firms and the S&P 400 mid-cap firms. For this, a distinction was made between the pre- and post-covid outbreak period. This showed that S&P 500 healthcare companies received relatively more attention after the outbreak, while the S&P 400 companies received less attention. One possible explanation may stem from the fact that the directly related companies identified later in the study are largely made up of S&P 500 companies. Logically they received more attention because of this direct link. In addition, existing literature shows that large companies generally receive more attention. This may also have impacted the attention that S&P 400 companies received.

Subsequently, an OLS and two fixed effects regressions were performed and compared for the period before Covid-19 (hypothesis 3a and 3b) and for the period during Covid-19 (4a and 4b). These regressions include all companies, both S&P 500 and S&P 400 healthcare firms. The focus of hypothesis 3 was on the short-term (first 1-2 weeks) and long-term (after reversal point) effect. Subsequently, in hypothesis 4, the same regression was performed for the period after the Covid-19 outbreak and assessed whether these coefficients were more positive. First, hypotheses 3a and 3b can be confirmed, since the reversal point occurs after 1 to 2 weeks in the pre-Covid period. The positive effect turns into a negative effect after this point. The same phenomenon was observed for the period after the Covid outbreak, although the exact point of the reversal point could differ. In addition, the results showed that the effect of investor attention on stock returns is more positive during Covid-19 in the short term (before the reversal point). However, in the long term, the effect of investor attention on stock returns during Covid-19 is more negative compared to the pre-Covid period.

Thereafter, a distinction is made between companies that are directly related to Covid-19 and companies that are not directly related. First, on the basis of the descriptive statistics of the SGSV

data, it is established that directly related companies received relatively more attention in the period after the outbreak. Relatively speaking, unrelated companies received less attention compared to the pre-Covid period. Subsequently, it is determined by means of the same regressions analysis as above that the effect in the long term (after reversal point) is more positive for the related companies. In contrast, the effect of investor attention in the short term is more negative for related companies. Finally, a comparison is made for the directly related companies between the period before and after the Covid-19 outbreak. On the basis of a comparison of two similar panel data regressions, it is determined that the effect in both short (1-2 weeks) and long (after reversal point) term is not unambiguous. Again, it is not possible to draw an unequivocal conclusion about the short term. However, it should be mentioned that the significance of the directly related companies has a share in the conclusion. Due to the lower number of companies (35 vs. 60), and therefore the lower number of observations, the results are more difficult to interpret for this group. Unfortunately, it seems slightly impossible to equal both numbers when your sample selection is based on an index like the S&P.

Then, the healthcare companies are divided into 8 groups based on firm size (4 quantiles) and relatedness to Covid-19 (related vs not related). Each quantile (firm size) can be combined with both related and unrelated companies, which allows eight combinations to be made. An analysis of the descriptive statistics during the SGSV data over the Covid-19 period shows that related companies in quantile 4 (largest firm size) actually received the most attention. Subsequently, the effect of investor attention during Covid-19 was measured using the combinations of firm, subsector, and week fixed effects regressions for 4 groups. Therefore, firm sizes from quantile 3 and 4 (largest firms) were merged together, as well as the firm sizes from quantiles 1 and 2 (smallest firms). This leaves two options in terms of firm size, the same number as the number of options in terms of relatedness. Therefore, four different combinations of firm size and relatedness can be created. The results showed that the groups including the largest firms (Q3+Q4) experienced the most positive effect of search volumes on the stock return. However, no definite conclusion can be drawn because for each lag, a number of SGSV outcomes are not significant. A possible reason is the fact that few observations remained per regression because the companies were subdivided into 4 groups. In addition, only the period after Covid-19 is included in the regressions for hypothesis 6a and 6b.

Finally, a pooled OLS regression was performed to gain insight into the different subsectors that make up the public healthcare sector. Based on dummy variables of 9 subsectors, a conclusion could be drawn about the effect of the individual subsectors. The coefficients hardly differ from each other. However, the results of the period before Covid-19 showed more positive results than the period after the outbreak.

Overall, the research question can be answered on the basis of the above discussed conclusions. The central question of this research focuses on the effect of SGSV data on stock returns and the differences between the pre- and post-Covid outbreak periods. The results and conclusions show that the effect of investor attention on stock returns of healthcare companies from the S&P 500

and 400 mid-cap is more positive for the first 1 to 2 weeks (short term) and more negative for the weeks after the reversal point (long term). Therefore, it can be concluded that Covid-19 has a reinforcing effect on the relationship between Google Search Volumes (investor attention) and stock returns. Hypothesis 4a and 4b showed a more positive effect in the short term and a more negative effect in the longer term during Covid-19. However, considering only the directly related companies to Covid-19 (hypothesis 5c), it cannot be concluded that Covid-19 has a moderating effect. As mentioned earlier, this can be caused by the fact that these specific regressions contain less observations due to the selection of only related companies and one of two time periods. In contrast, it was found that firm size is positively correlated with the amount of investor attention and its effect on stock returns. In addition to the amplifying effect of Covid-19 on the effect of SGSV on stock returns of healthcare companies from the S&P (compared to the pre-Covid period), it has also been established in all regressions that a reversal point occurred after 1 to 2 weeks, As found in existing literature.

Table 23. Hypotheses rejected/accepted

H1: Investor attention Granger causes returns in S&P 500 and S&P 400 mid-cap healthcare companies.	Accepted
H2a: During Covid-19 (from March 2020), investor attention for healthcare stocks from the S&P 500 exceeds the investor attention in the previous period (2019-Feb 2020).	Accepted
H2b: During Covid-19 (from March 2020), investor attention for healthcare stocks from the S&P 400 mid-cap exceeds the investor attention in the previous period (2019-Feb 2020).	Rejected
H3a: Investor attention for a healthcare company (large- and mid-cap) has a positive effect on the stock return in the short term (first 1-2 weeks) during the pre-Covid period.	Accepted
H3b: Investor attention for a healthcare company (large- and mid-cap) has a negative effect on the stock return in the long term (after reversal point) during the pre-Covid period.	Accepted
H4a: The effect of investor attention for a healthcare company (large- and mid-cap) on the stock return is more positive in the short term (first 1-2 weeks) during the Covid-19 crisis compared to the pre-Covid period.	Accepted
H4b: The effect of investor attention for a healthcare company (large- and mid-cap) on the stock return is more positive in the long term (after reversal point) during the Covid-19 crisis compared to the pre-Covid period.	Rejected
H5a: During Covid-19 (from March 2020), healthcare companies directly related to Covid-19 receive more attention than healthcare companies not directly related to Covid-19.	Accepted
H5b: The effect of investor attention for a healthcare company on the stock return is higher for healthcare companies directly related to the Covid-19 crisis than for healthcare companies not directly related to Covid-19 in both the short and long term.	Rejected (Only long term accepted)
H5c: The effect of investor attention for directly related healthcare companies on the stock return is higher during Covid-19 than in the pre-Covid period in both the short and long term.	Rejected
H6a: Large healthcare companies (based on firm size) directly related to Covid-19 receive the most investor attention during Covid-19.	Accepted
H6b: The effect of investor attention on the stock returns is the highest for directly related companies with the largest firm size.	Rejected (only the firm size part)
H7a: The effect of investor attention on stock returns for healthcare companies differs across healthcare subsectors.	Accepted
H7b: The effect of investor attention on stock returns for healthcare subsectors are more positive during Covid-19 compared to the period before.	Rejected

References

Bank, M., Larch, M., & Peter, G. (2011). Google search volume and its influence on liquidity and returns of German stocks. Financial Markets and Portfolio Management, 25(3). https://doi.org/10.1007/s11408-011-0165-y

Barber, B. M., & Odean, T. (2008). All That Glitters: The Effect of Attention and News on the Buying Behavior of Individual and Institutional Investors. *Review of Financial Studies*, *21*(2), 785–818. https://doi.org/10.1093/rfs/hhm079

Ben-Rephael, A., Da, Z., & Israelsen, R. D. (2017). It Depends on Where You Search: Institutional Investor Attention and Underreaction to News. The Review of Financial Studies, 30(9), 3009-3047. https://doi.org/10.1093/rfs/hhx031

Bijl, L., Kringhaug, G., Molnár, P., & Sandvik, E. (2016). Google searches and stock returns. *International Review of Financial Analysis*, *45*, 150–156. https://doi.org/10.1016/j.irfa.2016.03.015

Blitz, D., Huisman, R., Swinkels, L., & van Vliet, P. (2019). Media Attention and the Volatility Effect. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3403466

Cao, H. H., Han, B., Hirshleifer, D., & Zhang, H. H. (2009). Fear of the Unknown: Familiarity and Economic Decisions*. Review of Finance, 15(1), 173–206. https://doi.org/10.1093/rof/rfp023

DA, Z., ENGELBERG, J., & GAO, P. (2011). In Search of Attention. *The Journal of Finance*, *66*(5), 1461–1499. https://doi.org/10.1111/j.1540-6261.2011.01679.x

Demers, E., Hendrikse, J., Joos, P., & Lev, B. I. (2020). ESG Didn't Immunize Stocks Against the Covid-19 Market Crash. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3675920

Dey, A. K., Hoque, G. T., Das, K. P., & Panovska, I. (2022). Impacts of COVID-19 local spread and Google search trend on the US stock market. Physica A: Statistical Mechanics and Its Applications, 589, 126423. https://doi.org/10.1016/j.physa.2021.126423

Dimpfl, T., & Jank, S. (2015). Can Internet Search Queries Help to Predict Stock Market Volatility? European Financial Management, 22(2), 171–192. https://doi.org/10.1111/eufm.12058

Docherty, P., Chan, H., & Easton, S. (2011). Asset tangibility, industry representation and the cross section of equity returns. Australian Journal of Management, 36(1), 75–87. https://doi.org/10.1177/0312896211399200

He, P., Sun, Y., Zhang, Y., & Li, T. (2020). COVID-19's Impact on Stock Prices Across Different Sectors—An Event Study Based on the Chinese Stock Market. Emerging Markets Finance and Trade, 56(10), 2198–2212. https://doi.org/10.1080/1540496x.2020.1785865

Jiang, B., Zhu, H., Zhang, J., Yan, C., & Shen, R. (2021). Investor Sentiment and Stock Returns During the COVID-19 Pandemic. *Frontiers in Psychology*, *12*. https://doi.org/10.3389/fpsyg.2021.708537

Kedron, M. (2020). Goodwill and Its Effect on Share Price of Manufacturing and Nonmanufacturing Companies. Acta Universitatis Agriculturae et Silviculturae Mendelianae Brunensis, 68(2), 373–381. https://doi.org/10.11118/actaun202068020373

Kim, N., Lučivjanská, K., Molnár, P., & Villa, R. (2019). Google searches and stock market activity: Evidence from Norway. Finance Research Letters, 28, 208–220. https://doi.org/10.1016/j.frl.2018.05.003

Iyke, B.N., & Ho, S.Y. (2021). Investor attention on COVID-19 and African stock returns. MethodsX, 8, 101195. https://doi.org/10.1016/j.mex.2020.101195

Mazur, M., Dang, M., & Vega, M. (2020). COVID-19 and March 2020 Stock Market Crash. Evidence from S&P1500. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.3586603

Nguyen, C. P., Schinckus, C., & Hong Nguyen, T. V. (2019). Google search and stock returns in emerging markets. Borsa Istanbul Review, 19(4), 288–296. https://doi.org/10.1016/j.bir.2019.07.001

Oncu, E. (2021). The Impact of COVID-19 on Health Sector Stock Returns. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.3982420

Padungsaksawasdi, C., & Treepongkaruna, S. (2021). Chasing for information during the COVID-19 panic: The role of Google search on global stock market. Cogent Economics & Finance, 9(1). https://doi.org/10.1080/23322039.2021.1930669

Preis, T., Moat, H. S., & Stanley, H. E. (2013). Quantifying Trading Behavior in Financial Markets Using Google Trends. *Scientific Reports*, *3*(1). https://doi.org/10.1038/srep01684

Reilly, R.F. (2013). Reasons to Conduct the Health Care Intangible Asset Valuation. https://willamette.com/insights_journal/13/spring_2013_11.pdf

Rider, E. A., Comeau, M., Truog, R. D., Boyer, K., & Meyer, E. C. (2018). Identifying intangible assets in interprofessional healthcare organizations: feasibility of an asset inventory. Journal of Interprofessional Care, 33(5), 583–586. https://doi.org/10.1080/13561820.2018.1544118

Roberts, M.R., & Whited, T.M. (2011). Endogeneity in Empirical Corporate Finance. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.1748604

Sabbaghi, O. (2022). The impact of news on the volatility of ESG firms. Global Finance Journal, 51, 100570. https://doi.org/10.1016/j.gfj.2020.100570

Sheinson, D. M., Wong, W. B., Solon, C. E., Cheng, M. M., Shah, A., Elsea, D., & Meng, Y. (2020). Estimated Impact of Public and Private Sector COVID-19 Diagnostics and Treatments on US Healthcare Resource Utilization. Advances in Therapy, 38(2), 1212–1226. https://doi.org/10.1007/s12325-020-01597-3

Smales, L. A. (2020). Investor attention and the response of US stock market sectors to the COVID-19 crisis. *Review of Behavioral Finance*, *13*(1), 20–39. https://doi.org/10.1108/rbf-06-2020-0138

Smales, L. (2021). Investor attention and global market returns during the COVID-19 crisis. International Review of Financial Analysis, 73, 101616. https://doi.org/10.1016/j.irfa.2020.101616

Strasser, S., Stauber, C., Shrivastava, R., Riley, P., & O'Quin, K. (2021). Collective insights of public-private partnership impacts and sustainability: A qualitative analysis. PLOS ONE, 16(7), e0254495. https://doi.org/10.1371/journal.pone.0254495

Swamy, V., Dharani, M., & Takeda, F. (2019). Investor Attention and Google Search Volume Index: Evidence of an Emerging Market Using Quantile Regression Analysis. Research in International Business and Finance, 50, 1-17. https://doi.org/10.1016/j.ribaf.2019.04.010

Vasileiou, E. (2021). Explaining stock markets' performance during the COVID-19 crisis: Could Google searches be a significant behavioral indicator? Intelligent Systems in Accounting, Finance and Management, 28(3), 173–181. https://doi.org/10.1002/isaf.1499

Vlastakis, N., & Markellos, R.N. (2012). Information Demand and Stock Market Volatility. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.1558434

Vozlyublennaia, N. (2014). Investor attention, index performance, and return predictability. Journal of Banking & Finance, 41, 17–35. https://doi.org/10.1016/j.jbankfin.2013.12.010

Vrontis, D., el Chaarani, H., el Abiad, Z., el Nemar, S., & Yassine Haddad, A. (2021). Managerial innovative capabilities, competitive advantage and performance of healthcare sector during Covid-19 pandemic period. Foresight, 24(3/4), 504–526. https://doi.org/10.1108/fs-02-2021-0045

Wan, D., Xue, R., Linnenluecke, M., Tian, J., & Shan, Y. (2021). The impact of investor attention during COVID-19 on investment in clean energy versus fossil fuel firms. *Finance Research Letters*, *43*, 101955. https://doi.org/10.1016/j.frl.2021.101955

Xu, L., Chen, J., Zhang, X., & Zhao, J. (2020). COVID-19, public attention and the stock market. *Accounting & Finance*, *61*(3), 4741–4756. https://doi.org/10.1111/acfi.12734

Ying, Q., Kong, D., & Luo, D. (2015). Investor Attention, Institutional Ownership, and Stock Return: Empirical Evidence from China. Emerging Markets Finance and Trade, 51(3), 672-685. https://doi.org/10.1080/1540496x.2015.1046339

Yoshinaga, C., & Rocco, F. (2020). Investor Attention: Can Google Search Volumes Predict Stock Returns? *Brazilian Business Review*, *17*(5), 523–539. https://doi.org/10.15728/bbr.2020.17.5.3

Zhang, D., Hu, M., & Ji, Q. (2020). Financial markets under the global pandemic of COVID-19. *Finance Research Letters*, *36*, 101528. https://doi.org/10.1016/j.frl.2020.101528

Zhang, W., Shen, D., Zhang, Y., & Xiong, X. (2013). Open source information, investor attention, and asset pricing. Economic Modelling, 33, 613–619. https://doi.org/10.1016/j.econmod.2013.03.018

Appendix

Appendix A1: List of selected healthcare stocks from S&P 500

Table 24. Selected healthcare companies from S&P 500

Company Name	Company Ticker
Abbott Laboratories	**ABT**
AbbVie Inc.	ABBV
ABIOMED Inc.	**ABMD**
Agilent Technologies Inc.	**A**
Align Technology Inc.	ALGN
AmerisourceBergen Corporation	ABC
Amgen Inc.	AMGN
Anthem Inc.	ANTM
Baxter International Inc.	BAX
Becton Dickinson and Company	BDX
Biogen Inc.	BIIB
Bio-Rad Laboratories Inc. Class A	BIO
Bio-Techne Corporation	TECH
Boston Scientific Corporation	BSX
Bristol-Myers Squibb Company	BMY
Cardinal Health Inc.	CAH
Catalent Inc	CTLT
Centene Corporation	CNC
Cerner Corporation	CERN
Charles River Laboratories International Inc.	CRL
Cigna Corporation	CI
Cooper Companies Inc.	COO
CVS Health Corporation	CVS
Danaher Corporation	DHR
DaVita Inc.	DVA
DENTSPLY SIRONA Inc.	XRAY
DexCom Inc.	DXCM
Edwards Lifesciences Corporation	EW
Eli Lilly and Company	LLY
Gilead Sciences Inc.	GILD
HCA Healthcare Inc	HCA
Henry Schein Inc.	HSIC
Hologic Inc.	HOLX
Humana Inc.	HUM
IDEXX Laboratories Inc.	IDXX
Illumina Inc.	ILMN
Incyte Corporation	INCY
Intuitive Surgical Inc.	ISRG
IQVIA Holdings Inc	IQV

Johnson & Johnson	JNJ
Laboratory Corporation of America Holdings	LH
McKesson Corporation	MCK
Medtronic Plc	MDT
Merck & Co. Inc.	MRK
Mettler-Toledo International Inc.	MTD
Moderna Inc.	MRNA
Molina Healthcare Inc.	MOH
PerkinElmer Inc.	PKI
Pfizer Inc.	PFE
Quest Diagnostics Incorporated	DGX
Regeneron Pharmaceuticals Inc.	REGN
ResMed Inc.	RMD
STERIS Plc	STE
Stryker Corporation	SYK
Teleflex Incorporated	TFX
Thermo Fisher Scientific Inc.	TMO
UnitedHealth Group Incorporated	UNH
Universal Health Services Inc. Class B	UHS
Vertex Pharmaceuticals Incorporated	VRTX
Waters Corporation	WAT
West Pharmaceutical Services Inc.	WST
Zimmer Biomet Holdings Inc.	ZBH
Zoetis Inc. Class A	ZTS

Appendix A2: List of selected healthcare stocks from S&P 400 mid-cap

Table 25. Selected healthcare companies from S&P 400

Company Name	Company Ticker
Acadia Healthcare	ACHC
Amedisys	AMED
Arrowhead Pharmaceuticals	ARWR
Bruker	BRKR
Chemed Corp.	CHE
Encompass Health	EHC
Exelixis	EXEL
Globus Medical	GMED
Haemonetics	HAE
Halozyme	HALO
HealthEquity	HQY
Integra Lifesciences Holdings	IART
ICU Medical	ICUI
Jazz Pharmaceuticals	JAZZ
LHC Group	LHCG
LivaNova	LIVN
Masimo	MASI
Medpace	MEDP
Neurocrine Biosciences	NBIX
Neogen	NEOG
NuVasive	NUVA
Patterson Companies	PDCO
Penumbra	PEN
Perrigo Company Plc	PRGO
Quidel	QDEL
R1 RCM	RCM
Repligen	RGEN
STAAR Surgical	STAA
Syneos Health	SYNH
Tenet Health	THC
Tandem Diabetes Care	TNDM
United Therapeutics	UTHR

Appendix B: Descriptive Statistics Control Variables (adjusted)

Table 26. Descriptive Statistics Control Variables

Table 26 includes the adjusted descriptive statistics of the control variables. Therefore, the number of observations, mean, median, minimum, maximum, and standard deviation are shown per year.

	N	Mean	Median	Min	Max	St. Dev
Tobin's Q 2019	5035	3.578	2.654	1.059	13.083	2.466
Tobin's Q 2020	4940	3.998	2.883	1.026	19.340	3.158
Tobin's Q 2021	1615	4.155	2.594	1.005	23.563	3.569
Firm size 2019	5035	3.894	3.832	2.316	5.347	.711
Firm size 2020	4940	3.966	3.880	2.411	5.363	.684
Firm size 2021	1615	4.028	3.983	2.539	5.367	.668
RoE 2019	5035	0.324	.026	-.078	.159	.033
RoE 2020	4940	.017	.023	-.336	.219	.068
RoE 2021	1615	.038	.030	-.042	.188	.037
Leverage 2019	5035	.179	.614	-39.363	6.062	4.642
Leverage 2020	4940	1.294	.634	-39.363	34.946	6.622
Leverage 2021	1615	1.175	0.617	-39.363	34.946	6.020
MTB 2019	5035	6.057	4.546	-67.671	125.761	17.956
MTB 2020	4940	10.989	5.709	-67,671	125.761	22.547
MTB 2021	1615	9.982	5.073	-67.671	125.761	21.086
Goodwill 2019	5035	6.930	7.706	0	11.287	2.759
Goodwill 2020	4940	7.054	7.653	0	11.284	2.741
Goodwill 2021	1615	7.214	7.826	0	11.279	2.728
Intangible assets	5035	7.398	7.931	0	11.634	2.744
Intangible assets	4940	7.526	7.920	0	11.661	2.706
Intangible assets	1615	7.664	8.185	0	11.593	2.720
CPI 2019	5035	255.650	255.9	252.47	258.263	1.709
CPI 2020	4940	258.8446	258.845	255.944	261.564	1.677
CPI 2021	1615	264.367	265.028	262.2	266.727	1.671
Cap ex 2019	5035	5.132	5.043	1.396	8.333	1.559
Cap ex 2020	4940	5.201	5.233	.918	8.452	1.551
Cap ex 2021	1615	5.330	5.500	.376	8.400	1.578

Appendix C: Hausman results

Table 27. Hausman Test (SGSV by Ticker)

Table 27 shows the results of the Hausman test based on the SGSV by Ticker measure, which indicates whether a fixed effects or random effects model is preferred.

Test	Value
Chi Square	37.36
P-value	.0011

Table 28. Hausman Test (SGSV by Company Name)

Table 28 shows the results of the Hausman test based on the SGSV by Company Name measure, which indicates whether a fixed effects or random effects model is preferred.

Test	Value
Chi Square	38.05
P-value	.0009

The results of the two other SGSV measures show the same results. Therefore, the null hypothesis can be rejected. The null hypothesis states that the random effects model is preferred. Given the low p-values (.0009 < .05 and .0011 < .05) the null hypothesis will be rejected and the preferred model is the fixed effects model in all cases.

Appendix D: Directly vs not directly related firms (S&P 500 + S&P 400 mid-cap)

Table 29. Related Companies

Company Name	Company Ticker	Relatedness
Abbott Laboratories	ABT	Vaccine production
Abiomed Inc.	ABMD	Heart pump for Covid-19 patients
Agilent Technology Inc.	A	Covid-19 test
Amgen Inc.	AMGN	ACTIV, collaboration with Eli for anti-body drug
AmerisourceBergen Corp.	ABC	Vaccine distribution
Baxter International Inc.	BAX	Filling & Packaging
Becton Dickinson and Company	BDX	Covid-19 test
Bio-Rad Laboratories Inc.	BIO	Covid-19 test
Bio-Techne Corporation	TECH	Covid-19 test
Cardinal Health Inc.	CAH	Covid-19 test and other support
Catalent	CTLT	Vaccine development, leader
Charles River Laboratories International Inc.	CRL	Covid-19 test
CVS Health Corporation	CVS	Test locations
Danaher	DHR	Vaccine development and Covid-19 tests
Eli Lilly and Companies	LLY	Antibody drugs, ACTIV
Henry Schein	HSIC	Covid-19 test
Hologic Inc.	HOLX	Covid-19 test
IDEXX Laboratories Inc.	IDXX	Covid-19 test
Illumina Inc.	ILMN	Covid-19 test
Johnson & Johnson	JNJ	Vaccine production
Laboratory Corporation of America Holdings	LH	Covid-19 test (first for public)
McKesson Corporation	MCK	Distribution, Vaccinations, Packaging
Merck & Co. Inc	MRK	ACTIV, Drug development
Moderna	MRNA	Vaccine production
PerkinElmer Inc	PKI	Covid-19 test

Pfizer Inc.	PFE	Vaccine production
Quest Diagnostics Inc.	DGX	Covid-19 test
Regeneron Pharmaceuticals Inc.	REGN	Antibody cocktail
Thermo Fisher Scientific	TMO	Covid-19 test and supporting vaccine development
Waters Corporation	WAT	Covid-19 test
West Pharmaceutical Services Inc.	WST	Vaccine Distribution, and more
Zoetis Inc.	ZTS	Vaccine production (animals)
Neogen	NEOG	Wastewater detection screen
Quidel	QDEL	Covid-19 test
Repligen	RGEN	Production of part of vaccine

Appendix E: Optimal lag determination and Granger Causality tests

First, the optimal number of lags will be determined based on SGSV by Ticker and SGSS by Company Name. Based on the optimal number of lags, the Granger Causality tests will be performed to conclude on the direction of causality. The results in Appendix F serve as a robustness check on the SGSV Combined measure.

Table 30. Optimal Lag Determination (SGSV by Ticker)

Table 30 includes the results to determine the optimal number of lags to perform the Granger Causality tests, based on the SGSV by ticker measure. The lowest values of MBIC, MAIC, and MQIC determine the number of lags jointly. In addition, these lowest values are indicated by *.

Lag	CD	J	J Value	MBIC	MAIC	MQIC
1	-.167	52.001	.000	-96.356*	20.002	-19.264
2	-.165	27.066	.008	-84.203	3.065	-26.383*
3	-1.220	14.425	.071	-59.754	-1.575	-21.208
4	-4.114	3.983	.408	-33.106	-4.017*	-13.833

Table 31. Optimal Lag Determination (SGSV by Company Name)

Table 31 includes the results to determine the optimal number of lags to perform the Granger Causality tests, based on the SGSV by company name measure. The lowest values of MBIC, MAIC, and MQIC determine the number of lags jointly. In addition, these lowest values are indicated by *.

Lag	CD	J	J Value	MBIC	MAIC	MQIC
1	-.297	96.396	.000	-51.962	64.396	25.131
2	-.478	25.426	.013	-85.842*	1.427	-17.921
3	-1.627	17.712	.024	-56.467	1.712	--28.023*
4	-2.081	4.738	.315	-31.351	-3.261*	-13.078

The optimal number of lags for SGSV by Ticker data is based on an average. Therefore, it is decided to take 2 lags as optimal. For the SGSV by Company Name data, the same principle has been applied for the measure SGSV by Company Name. Therefore, the optimal number of lags is 3. The Granger Causality tests will be performed with a maximum of 3 lags and the results can be found in the tables below.

Table 32. Granger Causality tests (SGSV by Ticker)

Table 32 shows the results of the Granger Causality tests based on the SGSV by Ticker measure. Therefore, a Chi-square test is performed. The direction of the causality between investor attention (SGSV) and stock returns (Weekly Average) can be determined. ***, **, and *, indicate the 1%, 5%, and 10% significance levels, respectively.

Dependent var.	Excluded	Chi-sq	df	p-value
Weekly Average	SGSV by Ticker	12.087	2	.002***
	ALL	12.087	2	.002***
SGSV by Ticker	Weekly Average	5.150	2	.076*
	All	5.150	2	.076*

Table 33. Granger Causality tests (SGSV by Company Name)

Table 33 shows the results of the Granger Causality tests based on the SGSV by Company Name measure. Therefore, a Chi-square test is performed. The direction of the causality between investor attention (SGSV) and stock returns (Weekly Average) can be determined. ***, **, and *, indicate the 1%, 5%, and 10% significance levels, respectively.

Dependent var.	Excluded	Chi-sq	df	p-value
Weekly Average	SGSV by CN	25.179	3	.000***
	ALL	25.179	3	.000***
SGSV by CN	Weekly Average	6.708	3	.082*
	All	6.708	3	.082*

The Granger Causality tests show that the investor attention Granger-cause stock returns at 1% significance for both the SGSV by ticker as SGSV by company name measure. Stock returns only Granger-cause investor attention at a 10% level. The results are similar to the combined SGSV data.

Appendix F: Optimal lag determination panel data

In Table 34 the results regarding the lag determination can be found per SGSV measure. The control variables are also included, but not fully written out. The significance of the SGSV data is the most important for this determination. In addition, the lags of the dependent variable, weekly average stock return, are included in the standard panel data regression, as are week fixed effects. The only deviation is the fact that the first lag of SGSV by company name is not significant. However, weeks 1-4 are significant, so the conclusion is the same as with the other measures.

Table 34. Optimal lag determination panel data

This table presents the determination of the optimal number of lags for the regressions performed in this research. Therefore, for each SGSV measure a regression analysis is performed. The regressions include 6 lags of the dependent variable, 6 lags of the independent variable, and 6 lags of the control variables. ***, **, and *, indicate the 1%, 5%, and 10% significance levels, respectively.

	SGSV by Ticker		**SGSV by Company Name**		**SGSV Combined**	
Weekly Average	**Coefficient**	**Std. error.**	**Coefficient**	**Std. error.**	**Coefficient**	**Std. error.**
WA_L1	-.1031***	.0010	-.1035***	.0010	-.1041***	.0010
WA_L2	.0238***	.0010	.0227**	.0010	.0230**	.0010
WA_L3	-.0382***	.0010	-.0398***	.0010	-.0391***	.0010
WA_L4	-.0784***	.0010	-.0807***	.0010	-.0793***	.0010
WA_L5	-.0890***	.0010	-.0906***	.0010	-.0900***	.0010
WA_L6	.0010	.0010	.0010	.0010	.0014	.0010
Lag1	.0003*	.0001	.0002	.0001	.0003**	.0002
Lag2	-.0005***	.0001	-.0003***	.0001	-.0007***	.0002
Lag3	-.0003**	.0001	-.0003**	.0001	-.0004**	.0002
Lag4	-.0002*	.0001	-.0004***	.0001	-.0005***	.0002
Lag5	.0002	.0001	.0002	.0001	.0003	.0002
Lag6	.0001	.0001	-.0001*	.0001	-.0001	.0002
Controls	Yes		Yes		Yes	
N	11020		11020		11020	
Firms	95		95		95	
Lags	6		6		6	
F-test	7.02***		7.08***		7.28***	
R2	.0359		.0353		.0368	

Appendix G: Ticker and Company Name Results Hypothesis 3

Table 35(1). Effect of investor attention by SGSV by Ticker (pre-Covid)

This table presents three regressions regarding the effect of investor attention during the pre-Covid period, measured by SGSV by ticker data. Therefore, the sample period is January 2019 to February 2020. The regressions showed in this table are an OLS, subsector + week fixed effects, and firm + week fixed effects regression. Moreover, the table includes 4 lags of the SGSV data, a lag of the dependent variable, and a lag of the control variables. The F-test indicates the significance of the entire model and the R2-adjusted is, compared to the R2, adjusted for the number of predictor variables in the model. ***, **, and *, indicate the 1%, 5%, and 10% significance levels, respectively.

Weekly Average	**OLS**		**SubSector + Week FE**		**Firm + Week FE**	
	Coefficient	**St. Error**	**Coefficient**	**St. Error**	**Coefficient**	**St. Error**
WA_L1	-.0711***	.0144	-.0262*	.014	-.0361	.0141
Lag1	.0001	.0002	.0001	.0001	.0001	.0001
Lag2	-.0002	.0002	-.0002*	.0001	-.0002*	.0001
Lag3	-.0002	.0002	-.0001	.0001	-.0001	.0001
Lag4	-.0003*	.0002	-.0001*	.0001	-.0001*	.0001
RoE_L1	.0008	.0038	-.0035	.0033	-.0038	.0057
MBR_L1	.0000	.0000	.0000	.0000	.0000	.0000
Leverage_L1	.0000	.0000	.0000	.0000	-.0001	.0001
Intangibles_L1	.0004	.0003	.0004	.0003	.002	.0017
Goodwill_L1	-.0004	.0003	-.0004	.0003	-.0039	.0018
Firm Size_L1	-.0009*	.0005	-.0005	.0006	.0063	.0045
Cap. Ex._L1	.0001	.0002	.0001	.0002	.0016*	.0010
Tobin's Q_L1	.0000	.0001	.0000	.0001	.0002	.0003
CPI_L1	-.0002***	.0001	-.0038***	.0002	-.0039***	.0002
Constant	.0582***	.0194	.9624***	.0461	.9565***	.0466
SubSector FE	No		Yes		No	
Firm FE	No		No		Yes	
Week FE	No		Yes		Yes	
N	5415		5415		5415	
Firms	95		95		95	
Lags	4		4		4	
F-Test	3.8704***		35.8263***		35.8769***	
R2-adjusted	.0074		.3066		.2993	

Table 35(2). Effect of investor attention by SGSV by Company Name (pre-Covid)

This table presents three regressions regarding the effect of investor attention during the pre-Covid period, measured by SGSV by company name data. Therefore, the sample period is January 2019 to February 2020. The regressions showed in this table are an OLS, subsector + week fixed effects, and firm + week fixed effects regression. Moreover, the table includes 4 lags of the SGSV data, a lag of the dependent variable, and a lag of the control variables. The F-test indicates the significance of the entire model and the R2-adjusted is, compared to the R2, adjusted for the number of predictor variables in the model. ***, **, and *, indicate the 1%, 5%, and 10% significance levels, respectively.

Weekly Average	**OLS**		**SubSector + Week FE**		**Firm + Week FE**	
	Coefficient	**St. Error**	**Coefficient**	**St. Error**	**Coefficient**	**St. Error**
WA_L1	-.0703***	.0144	-.0260*	.0140	-.0362**	.0141
Lag1	.0003*	.0002	.0002	.0001	.0002	.0001
Lag2	.0003*	.0002	.0003**	.0001	.0003**	.0001
Lag3	-.0004**	.0002	-.0001	.0001	-.0001	.0001
Lag4	-.0004***	.0002	-.0003*	.0001	-.0002	.0001
RoE_L1	.0004	.0038	-.0037	.0033	-.0026	.0057
MBR_L1	.0000	.0000	.0000	.0000	.0000	.0000
Leverage_L1	.0000	.0000	.0000	.0000	-.0001	.0001
Intangibles_L1	.0003	.0003	.0003	.0003	.0019	.0017
Goodwill_L1	-.0003	.0003	-.0003	.0003	-.0038**	.0018
Firm Size_L1	-.0008	.0005	-.0005	.0006	.006	.0045
Cap. Ex._L1	.0002	.0002	.0001	.0002	.0015	.001
Tobin's Q_L1	.0000	.0001	.0000	.0001	.0002	.0003
CPI_L1	-.0002***	.0001	-.0038***	.0002	-.0039***	.0002
Constant	.0555***	.0193	.9684***	.0455	.9560***	.0465
SubSector FE	No		Yes		No	
Firm FE	No		No		Yes	
Week FE	No		Yes		Yes	
N	5415		5415		5415	
Firms	95		95		95	
Lags	4		4		4	
F-Test	4.3238***		35.9165***		35.9827***	
R2-adjusted	.0085		.3072		.3210	

Appendix H: Ticker and Company Name Results Hypothesis 4

Table 36(1). Effect of investor attention by SGSV by Ticker (post-Covid)

This table presents three regressions regarding the effect of investor attention during the Covid-19 period, measured by SGSV by ticker data. Therefore, the sample period is February 2020 to April 2021. The regressions showed in this table are an OLS, subsector + week fixed effects, and firm + week fixed effects regression. Moreover, the table includes 4 lags of the SGSV data, a lag of the dependent variable, and a lag of the control variables. The F-test indicates the significance of the entire model and the R2-adjusted is, compared to the R2, adjusted for the number of predictor variables in the model. ***, **, and *, indicate the 1%, 5%, and 10% significance levels, respectively.

Weekly Average	**OLS**		**SubSector + Week FE**		**Firm + Week FE**	
	Coefficient	**St. Error**	**Coefficient**	**St. Error**	**Coefficient**	**St. Error**
WA_L1	-.1152***	.0127	-.1062***	.013	-.1180***	.0131
Lag1	.0004**	.0002	.0008***	.0002	.0008***	.0002
Lag2	-.0008***	.0002	-.0008***	.0002	-.0008***	.0002
Lag3	-.0003	.0002	-.0003*	.0002	-.0004**	.0002
Lag4	-.0002	.0002	-.0002	.0002	-.0002	.0002
RoE_L1	-.0003	.0030	-.0006	.0029	.0043	.0051
MBR_L1	.0000	.0000	.0000	.0000	.0000	.0000
Leverage_L1	.0000	.0001	.0000	.0000	.0000	.0001
Intangibles_L1	-.0007*	.0004	-.0005*	.0003	-.0021	.0021
Goodwill_L1	.0006	.0004	.0004	.0003	.0039*	.0023
Firm Size_L1	.0008	.0007	.0010	.0006	-.0025	.0044
Cap. Ex._L1	-.0003	.0002	-.0004	.0002	-.0015*	.0008
Tobin's Q_L1	.0001	.0001	.0001*	.0001	.0003	.0003
CPI_L1	-.0002***	.0001	-.0008***	.0002	-.0008***	.0002
Constant	.0587***	.0157	.2139***	.0536	.2145***	.0542
SubSector FE	No		Yes		No	
Firm FE	No		No		Yes	
Week FE	No		Yes		Yes	
N	5795		5795		5795	
Firms	95		95		95	
Lags	4		4		4	
F-Test	9.3103***		50.3596***		50.3609***	
R2-adjusted	.0197		.3828		.3772	

Table 36(2). Effect of investor attention by SGSV by Company Name (post-Covid)

This table presents three regressions regarding the effect of investor attention during the Covid-19 period, measured by SGSV by company name data. Therefore, the sample period is February 2020 to April 2021. The regressions showed in this table are an OLS, subsector + week fixed effects, and firm + week fixed effects regression. Moreover, the table includes 4 lags of the SGSV data, a lag of the dependent variable, and a lag of the control variables. The F-test indicates the significance of the entire model and the R2-adjusted is, compared to the R2, adjusted for the number of predictor variables in the model. ***, **, and *, indicate the 1%, 5%, and 10% significance levels, respectively.

Weekly Average	**OLS**		**SubSector + Week FE**		**Firm + Week FE**	
	Coefficient	**St. Error**	**Coefficient**	**St. Error**	**Coefficient**	**St. Error**
WA_L1	-.1169***	.0127	-.1077***	.0130	-.1193***	.0131
Lag1	.0003*	.0002	.0008***	.0002	.0008***	.0002
Lag2	-.0008***	.0002	-.0006***	.0002	-.0006***	.0002
Lag3	-.0003*	.0002	-.0005***	.0002	-.0005***	.0002
Lag4	-.0003*	.0002	-.0003*	.0002	-.0003*	.0002
RoE_L1	-.0001	.0030	-.0008	.0029	.0035	.0051
MBR_L1	.0000	.0000	.0000	.0000	.0000	.0000
Leverage_L1	.0000	.0001	.0000	.0000	-.0001	.0001
Intangibles_L1	-.0006*	.0004	-.0005	.0003	-.0019	.0021
Goodwill_L1	.0005	.0004	.0003	.0003	.0035	.0023
Firm Size_L1	.0006	.0007	.0009	.0006	-.0016	.0044
Cap. Ex._L1	-.0002	.0002	-.0004	.0002	-.0015*	.0008
Tobin's Q_L1	.0001	.0001	.0001*	.0001	.0003	.0003
CPI_L1	-.0002***	.0001	-.0008***	.0002	-.0008***	.0002
Constant	.0508***	.0157	.2145***	.0536	.2131***	.0541
SubSector FE	No		Yes		No	
Firm FE	No		No		Yes	
Week FE	No		Yes		Yes	
N	5795		5795		5795	
Firms	95		95		95	
Lags	4		4		4	
F-Test	9.9065		50.4634***		50.4386***	
R2-adjusted	.0211		.3833		.3776	

Appendix I: Ticker and Company Name Results Hypothesis 5b

Table 37(1). Effect of directly related companies (SGSV by Ticker)

This table presents three regressions regarding the effect of investor attention measured by SGSV by ticker data. It only includes firms that are considered as directly related to the Covid-19 pandemic. Therefore, the sample period is January 2019 to April 2021. The regressions showed in this table are an OLS, subsector + week fixed effects, and firm + week fixed effects regression. Moreover, the table includes 4 lags of the SGSV data, a lag of the dependent variable, and a lag of the control variables. The F-test indicates the significance of the entire model and the R2-adjusted is, compared to the R2, adjusted for the number of predictor variables in the model. ***, **, and *, indicate the 1%, 5%, and 10% significance levels, respectively.

Weekly Average	**OLS**		**SubSector + Week FE**		**Firm + Week FE**	
	Coefficient	**St. Error**	**Coefficient**	**St. Error**	**Coefficient**	**St. Error**
WA_L1	-.0505***	.0156	-.0459***	.0158	-.0564***	.0158
Lag1	-.0001	.0002	.0001	.0002	.0001	.0002
Lag2	-.0001	.0002	-.0004**	.0002	-.0004**	.0002
Lag3	-.0003	.0002	-.0002	.0002	-.0002	.0002
Lag4	.0001	.0002	-.0000	.0002	-.0000	.0002
RoE_L1	-.0008	.0030	-.001	.0027	-.0002	.004
MBR_L1	.0000	.0000	.0000	.0000	.0000	.0000
Leverage_L1	.0000	.0000	.0000	.0000	-.0001	.0001
Intangibles_L1	-.0005	.0009	-.0011	.0009	-.0034	.0031
Goodwill_L1	.0003	.0009	.001	.0009	.0031	.0032
Firm Size_L1	.0013*	.0007	.0017**	.0007	.0085***	.0025
Cap. Ex._L1	-.0006*	.0003	-.0009***	.0003	-.0018**	.0008
Tobin's Q_L1	.0001	.0001	.0001	.0001	.0004**	.0002
CPI_L1	.0000	.0001	-.0006	.0001	-.0007***	.0001
Constant	.0094	.0139	.1569***	.0377	.1541***	.0379
SubSector FE	No		Yes		No	
Firm FE	No		No		Yes	
Week FE	No		Yes		Yes	
N	4130		4130		4130	
Firms	35		35		35	
Lags	4		4		4	
F-Test	2.0070***		16.3183***		16.4537***	
R2-adjusted	.0034		.3246		.3236	

Table 37(2). Effect of companies not directly related (SGSV by Ticker)

This table presents three regressions regarding the effect of investor attention measured by SGSV by ticker data. It only includes firms that are considered as not directly related to the Covid-19 pandemic. Therefore, the sample period is January 2019 to April 2021. The regressions showed in this table are an OLS, subsector + week fixed effects, and firm + week fixed effects regression. Moreover, the table includes 4 lags of the SGSV data, a lag of the dependent variable, and a lag of the control variables. The F-test indicates the significance of the entire model and the R2-adjusted is, compared to the R2, adjusted for the number of predictor variables in the model. ***, **, and *, indicate the 1%, 5%, and 10% significance levels, respectively.

Weekly Average	**OLS**		**SubSector + Week FE**		**Firm + Week FE**	
	Coefficient	**St. Error**	**Coefficient**	**St. Error**	**Coefficient**	**St. Error**
WA_L1	-.1179***	.0118	-.0880***	.0119	-.0924***	.0120
Lag1	.0005***	.0002	.0006***	.0001	.0006***	.0001
Lag2	-.0007***	.0002	-.0005***	.0002	-.0005***	.0002
Lag3	-.0003*	.0002	-.0002*	.0002	-.0002*	.0002
Lag4	-.0003*	.0002	-.0002	.0001	-.0002	.0001
RoE_L1	-.0020	.0033	-.0004	.0031	.0005	.0038
MBR_L1	.0000	.0000	.0000	.0000	.0000	.0000
Leverage_L1	.0000	.0000	.0000	.0000	.0000	.0000
Intangibles_L1	.0001	.0003	.0003	.0003	.0004	.0009
Goodwill_L1	.0000	.0003	-.0002	.0002	.0004	.0009
Firm Size_L1	-.0005	.0006	-.0004	.0006	-.0038	.0032
Cap. Ex._L1	.0000	.0002	-.0001	.0002	.0001	.0005
Tobin's Q_L1	.0001*	.0001	.0002***	.0001	.0004**	.0002
CPI_L1	.0000	.0000	-.0004***	.0001	-.0003***	.0001
Constant	-.0053	.0123	.0929***	.0325	.0971***	.0329
SubSector FE	No		Yes		No	
Firm FE	No		No		Yes	
Week FE	No		Yes		Yes	
N	7080		7080		7080	
Firms	60		60		60	
Lags	4		4		4	
F-Test	10.0861***		36.4370***		36.3524***	
R2-adjusted	.0177		.3938		.3906	

Table 38(1). Effect of directly related companies (SGSV by Company Name)

This table presents three regressions regarding the effect of investor attention measured by SGSV by company name data. It only includes firms that are considered as directly related to the Covid-19 pandemic. Therefore, the sample period is January 2019 to April 2021. The regressions showed in this table are an OLS, subsector + week fixed effects, and firm + week fixed effects regression. Moreover, the table includes 4 lags of the SGSV data, a lag of the dependent variable, and a lag of the control variables. The F-test indicates the significance of the entire model and the R2-adjusted is, compared to the R2, adjusted for the number of predictor variables in the model. ***, **, and *, indicate the 1%, 5%, and 10% significance levels, respectively.

Weekly Average	**OLS**		**SubSector + Week FE**		**Firm + Week FE**	
	Coefficient	**St. Error**	**Coefficient**	**St. Error**	**Coefficient**	**St. Error**
WA_L1	-.0522***	.0156	-.0448***	.0158	-.0556***	.0158
Lag1	.0001	.0002	.0003**	.0002	.0003*	.0002
Lag2	-.0004*	.0002	-.0001*	.0002	-.0001*	.0002
Lag3	-.0002	.0002	-.0002**	.0002	-.0002**	.0002
Lag4	-.0001	.0002	-.0000	.0002	-.0000	.0002
RoE_L1	-.0004	.003	-.0010	.0028	-.0003	.0040
MBR_L1	.0000	.0000	.0000	.0000	.0000	.0000
Leverage_L1	.0000	.0000	.0000	.0000	.0000	.0001
Intangibles_L1	-.0004	.0009	-.0011	.0009	-.0037	.0031
Goodwill_L1	.0002	.0009	.0010	.0009	.0033	.0032
Firm Size_L1	.0013*	.0007	.0016**	.0007	.0084***	.0025
Cap. Ex._L1	-.0006*	.0003	-.0009***	.0003	-.0017**	.0008
Tobin's Q_L1	.0001	.0001	.0001	.0001	.0004**	.0002
CPI_L1	.0000	.0001	-.0006***	.0001	-.0007***	.0001
Constant	.0103	.0135	.1557***	.0378	.1532***	.0379
SubSector FE	No		Yes		No	
Firm FE	No		No		Yes	
Week FE	No		Yes		Yes	
N	4130		4130		4130	
Firms	35		35		35	
Lags	4		4		4	
F-Test	2.5341***		16.2935***		16.4328***	
R2-adjusted	.0052		.3242		.3233	

Table 38(2). Effect of companies not directly related (SGSV by Company Name)

This table presents three regressions regarding the effect of investor attention measured by SGSV by company name data. It only includes firms that are considered as not directly related to the Covid-19 pandemic. Therefore, the sample period is January 2019 to April 2021. The regressions showed in this table are an OLS, subsector + week fixed effects, and firm + week fixed effects regression. Moreover, the table includes 4 lags of the SGSV data, a lag of the dependent variable, and a lag of the control variables. The F-test indicates the significance of the entire model and the R2-adjusted is, compared to the R2, adjusted for the number of predictor variables in the model. ***, **, and *, indicate the 1%, 5%, and 10% significance levels, respectively.

Weekly Average	**OLS**		**SubSector + Week FE**		**Firm + Week FE**	
	Coefficient	**St. Error**	**Coefficient**	**St. Error**	**Coefficient**	**St. Error**
WA_L1	-.1188***	.0118	-.0890***	.0119	-.0934***	.0120
Lag1	.0003**	.0002	.0006***	.0001	.0006***	.0001
Lag2	-.0003	.0002	-.0003	.0001	-.0003	.0001
Lag3	-.0004**	.0002	-.0004	.0001	-.0004	.0001
Lag4	-.0005***	.0002	-.0004***	.0001	-.0004***	.0001
RoE_L1	-.0021	.0033	-.0004	.0031	.0006	.0038
MBR_L1	.0000	.0000	.0000	.0000	.0000	.0000
Leverage_L1	.0000	.0000	.0000	.0000	.0000	.0000
Intangibles_L1	.0001	.0003	.0003	.0003	.0005	.0009
Goodwill_L1	.0000	.0003	-.0002	.0002	.0003	.0009
Firm Size_L1	-.0005	.0006	-.0004	.0006	-.0035	.0032
Cap. Ex._L1	.0000	.0002	-.0001	.0002	.0001	.0006
Tobin's Q_L1	.0001	.0001	.0002**	.0001	.0004**	.0002
CPI_L1	.0000	.0000	-.0004***	.0001	-.0004***	.0001
Constant	-.0054	.0122	.0932***	.0326	.0968***	.0329
SubSector FE	No		Yes		No	
Firm FE	No		No		Yes	
Week FE	No		Yes		Yes	
N	7080		7080		7080	
Firms	60		60		60	
Lags	4		4		4	
F-Test	10.0454***		36.4357***		36.3518***	
R2-adjusted	.0176		.3937		.3906	