Beyond the Numbers: The Role of Attention in Stock Returns

Miriam Cruz

Dissertation written under the supervision of Professor Eva Schliephake

Dissertation submitted in partial fulfilment of requirements for the MSc in Finance, at the Universidade Católica Portuguesa, 6th January 2025.

Abstract

This paper investigates whether the Abnormal Google Search Volume Index (ASVI) predicts Abnormal Returns in the Nasdaq-100, using the company's names combined with the word 'stock' as a keyword. Analysing data from all Nasdaq-100 stocks between 2019 and 2023, with weekly frequency, I confirm that (1) the ASVI, as a proxy for investor attention, provides unique insights that the other proxies cannot explain, as in Da et al. (2011). Additionally, I find that (2) an increase in investor attention, captured by ASVI, predicts higher short-term abnormal results, with the strongest effects observed during periods of high market volatility, such as during and after the COVID-19 period. Lastly, I find that (3) an increase in ASVI has a stronger impact on the increase of Abnormal Return for technology stocks. These conclusions were mainly based on the CCEMG model, which is particularly suited for analysing macroeconomic shocks. In contrast, GMM cannot capture these effects without further adjustments.

Keywords: Abnormal Google Search Volume Index, Investor Attention, Abnormal Return, COVID-19 Pandemic, Technology Stocks

Title: Beyond the Numbers: The Role of Attention in Stock Returns

Author: Miriam Cruz

Resumo

Este artigo investiga se o Índice Anormal de Volume de Pesquisas do Google (ASVI) prevê Retornos Anormais no Nasdaq-100, utilizando como palavras-chave o nome das empresas em conjunto com o termo 'ação'. Analisando os dados de todas as ações do Nasdaq-100, entre 2019 e 2023, com frequência semanal, confirmo que (1) o ASVI, enquanto indicador da atenção dos investidores, fornece informações únicas que outros indicadores não conseguem explicar, tal como demonstrado em Da et al. (2011). Adicionalmente, concluo que (2) um aumento da atenção dos investidores, captado pelo ASVI, prevê Retornos Anormais de curto prazo mais elevados, com os efeitos mais acentuados observados durante períodos de elevada volatilidade do mercado, como durante e depois do período da COVID-19. Por fim, verifico que (3) um aumento no ASVI tem um impacto mais pronunciado no aumento dos Retornos Anormais das ações tecnológicas. Estas conclusões baseiam-se principalmente no modelo CCEMG, uma vez que este é o modelo mais adequado para analisar choques macroeconómicos. Em contraste, o modelo GMM não consegue captar totalmente estes efeitos sem ajustes adicionais.

Palavras-chave: Anormal Índice de Volume de Pesquisas do Google, Atenção dos Investidores, Retornos Anormais, Pandemia de COVID-19, Ações Tecnológicas

Título: Para Além dos Números: O papel da Atenção nos Retornos das Ações

Autora: Miriam Cruz

Acknowledgements

First, I would like to thank my Supervisor, Professor Eva Schliephake, for her guidance and valuable insights during these last four months.

I am deeply grateful to the Refinitiv team for their tireless support and dedication in ensuring I had access to the data necessary for my research.

I am also indebted to the friends I made during this master's for all the support during this challenging period, without their presence, it would have been much harder to have completed this master's. Since with this experience, it has proven once more that 'Alone we go faster, together we go further'.

I would also like to thank my colleagues at Siemens for their support during these demanding months, which allowed me to reconcile work with the writing of this thesis.

Finally, I would like to express my gratitude to my family, close friends, and boyfriend, for their unconditional support and understanding, and for encouraging me to persist through every challenge.

I also acknowledge the use of AI to brainstorm research ideas and as a copy editor to improve grammar and language precision, ensuring the quality of this thesis.

Table of Contents

List of Tables

List of Figures

List of Appendices

List of Abbreviations

AIA - Abnormal Institutional Investor Attention

ASVI – Abnormal Google Search Volume Index

BPS – Basis Points

CCEMG – Common Correlated Effects Mean Group

CRSP – Centre for Research in Security Prices

EMH – Efficient Market Hypothesis

FE – Fixed Effects

GMM – Generalized Method of Moments

HML – High Minus Low

IPO – Initial Public Offering

IVOL – Idiosyncratic Volatility

Mkt-Rf – Market Return minus the Risk-Free Rate

OLS – Ordinary Least Squares

RE – Random Effects

SMB – Small Minus Big

SVI – Google Search Volume Index

VAR – Vector Autoregression

VIF – Variance Inflation Factor

I. Introduction

The number of individuals investing has been growing over the past decade, peaking in 2021 (Wheat & Eckerd, 2024). This trend results, among other factors, from the greater accessibility to stock markets, facilitated by the rise of online brokerage platforms (Statista Research Department, 2023). However, the rise in accessibility did not correspond to an increase in financial literacy among retail investors (Angrisani et al., 2023), leaving retail investors at a disadvantage relative to institutional investors, who leverage sophisticated algorithms to achieve Abnormal Returns.

The Efficient Market Hypothesis (EMH) states that asset prices immediately incorporate all available information, making it nearly impossible to outperform the market consistently (Fama, 1970; Merton, 1987). However, EMH assumes that investors act rationally and optimally (Fama, 1970), which does not always align with reality since investors are influenced by cognitive biases and face limited resources, becoming overwhelmed by the volume of available metrics (Hirshleifer & Teoh, 2003; Peng, 2005). Recognising these limitations, I explore the potential to achieve Abnormal Returns by focusing on certain behavioural indicators, particularly those that measure investor attention because there is still a lack of evidence on the limits and applications of the Google Search Volume Index (SVI).

Moreover, there is also a gap in understanding how this proxy performs over prolonged and high-volatile periods, such as the COVID-19 pandemic, and in specific sectors, namely the technological. The COVID-19 pandemic period and the technological sector can be particularly interesting to study because both present higher volatility compared to other periods (Basuony et al., 2021; Kayani et al., 2024; Ullah et al., 2023) and other sectors (Moran, 2020), respectively. Additionally, according to Yahya et al. (2021), during the COVID-19 pandemic, investors had more time to pay attention to the markets and make more informed decisions. Therefore, this is worthwhile studying because Padungsaksawasdi et al. (2019) conclude that volatility attracts more attention due to uncertainty, and more attention ultimately leads to higher Abnormal Returns, as in Da et al. (2011). Understanding these effects can help not only retail investors but the entire field of finance to predict and analyse market dynamics.

This paper studies the ASVI, which is the abnormal value of SVI, a relative measure of popularity from Google Trends regarding a particular keyword, in this case, the company names and the word 'stock' as explained in Data. Google Trends is a free tool, that collects real-time, worldwide data from Google, the most well-known search engine, as seen in Figure I (Bianchi, 2024). This tool provides daily data access, eliminating reliance on delayed corporate data releases.

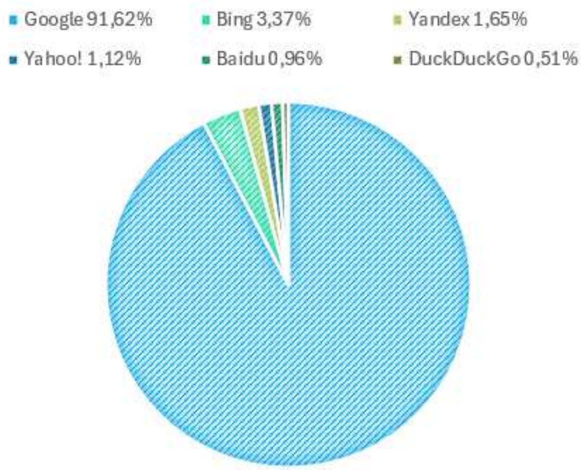

Figure I. Global search engine market share | Statista

Google is especially valuable to the financial field as its platform is well known for both institutional and retail investors, contrary to other platforms, like Bloomberg or Reuters, which are widely known among institutional investors but less accessible to the majority of retail investors.

Figure II demonstrates the specific example of Datadog stock, that until 19/09/2019, the Initial Public Offering (IPO) date, presented no relevant searches for the company, and from that point onwards the searches increased particularly in the IPO week, illustrating that SVI reflects real-time shifts in investor attention around major events, which shows the unique insights that SVI can provide (Google, 2024).

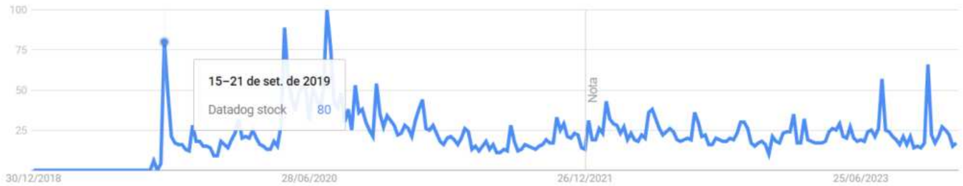

Figure II. Google Search Volume Index from Datadog (Keyword: Datadog stock)

Therefore, this thesis aims to study whether the Google Abnormal Search Volume Index (ASVI), as a proxy for investor attention, can predict Abnormal Returns in the Nasdaq-100, particularly in technology stocks and during periods of high volatility.

This study focuses on Abnormal Returns to isolate stock-specific effects from general market movements, with a particular emphasis on the Nasdaq-100. This index includes some of the most well-known firms worldwide that tend to attract significant retail investor attention, making it an ideal index for studying the impact of investor behaviour on stock performance. Additionally, analysing the COVID-19 period allows for valuable insights into the effects of uncertain market conditions.

Looking at the literature, Da et al. (2011) are the ones who introduce the Abnormal Google Search Volume Index as a proxy for investor attention. Nevertheless, the concept of investor attention first appears in the literature with Simon (1971), who argues that in a world rich in information, attention becomes a scarce resource, which means that attention needs to be efficiently allocated. This concept gains particular importance after 1970, as the literature moved from traditional finance models to behavioural finance models. Unlike the Efficient Market Hypothesis (EMH), which assumes all information is instantly reflected in the financial markets (Fama, 1970), behavioural finance models state that not all investors are fully rational (Thaler, 2005). Merton (1987) and Barber & Odean (2008) further build on this, emphasising that investors tend to buy stocks of which they are aware, which typically means attention-grabbing stocks, introducing a bias in their investments by not being rational.

Following the rise of behavioural finance models, several papers appear and attempt to measure investor attention, through indirect proxies, such as Trading Volume (Barber & Odean, 2008; Gervais et al., 2001; Preis et al., 2010), Advertising/Sales (Grullon et al., 2004; Lou, 2014), Analyst Coverage (Da et al., 2011), News Coverage from journals (Barber & Odean, 2008; Tetlock, 2007; Thompson II et al., 1987), News Coverage from digital media platforms (Das & Chen, 2007; Mondria et al., 2010; Siganos et al., 2014), and Abnormal Institutional Investor Attention (AIA), which uses news from Bloomberg terminals and allows the conclusion that institutional investors respond quicker to major news events (Ben-Rephael et al., 2017).

In May 2006, Google launched Google Trends, introducing a direct measure of investor attention, the Abnormal Search Volume Index (ASVI), that provides unique insights into investor attention, as explained in Hypothesis I. This tool allows the creation of many more studies with various purposes, such as, anticipating the widespread of a disease (Ginsberg et al., 2009), forecasting unemployment (Askitas & Zimmermann, 2009), and predicting automotive sales, home sales, and tourism (Varian & Choi, 2012). Other studies focus on trying to predict returns based on information not yet priced in the market and discovered that an increase in SVI predicts temporary higher returns with subsequent return reversal (Bank et al., 2011; Bijl et al., 2016; Da et al., 2011; Tan & Taş, 2019; Yoshinaga & Rocco, 2020). In this thesis, particularly in Hypothesis II, I also focus on trying to predict Abnormal Returns based on SVI but I do not focus on the price reversal aspect, as Ayala et al. (2024) and Ayaz et al.

(2021) who found out that an increase in the Google Search Volume Index is associated with positive returns.

Nevertheless, other studies challenge these conclusions, such as Lobão et al. (2017), who find weak evidence of Abnormal Returns in the following month due to increases in the monthly search volume. However, this discrepancy can be justified by the fact that Lobão et al.'s paper presents a monthly frequency, in contrast to the weekly frequency used in most other studies. Therefore, the short-term increases in Abnormal Returns may be offset by the return reversal, as supported by other papers' findings.

The impact of different market conditions, in turn, is the subject of only a few studies. Therefore, distinctively from Da et al. (2011), this thesis adds by analysing each period of the COVID-19 pandemic in Hypothesis II, which can be very interesting since this period has unique characteristics, such as high volatility (Basuony et al., 2021; Kayani et al., 2024; Ullah et al., 2023) and more time for investors to pay attention to the markets and make better decisions (Yahya et al., 2021). Some studies in the literature explore the impact of SVI during the COVID-19 period, but they all have a different focus from this thesis since I use keywords related to each firm, instead of using keywords related to the macro conditions, such as 'recession', 'COVID', 'bankruptcy', 'vaccines', 'depression' and others (Cevik et al., 2022; Chundakkadan & Nedumparambil, 2022; Da et al., 2015; Smales, 2020; Smales, 2021).

As in Da et al. (2011), this thesis focuses on the impact of ASVI on Abnormal Returns, but I also add the focus on the difference between technology and non-technology stocks, which will be explored in Hypothesis III. The focus on the technology sector can be interesting because the technology sector presents high innovation, growth potential and volatility, which tend to attract more investor attention (Cutter, 2024; Moran, 2020).

To predict Abnormal Returns using the ASVI, this thesis addresses the following hypotheses:

Hypothesis I: The Abnormal Google Search Volume Index differs from other investor attention proxies.

This hypothesis tests whether ASVI provides unique insights compared to traditional proxies such as Analyst Coverage, Advertising/Sales and News Coverage. I expect ASVI to outperform these traditional proxies in predicting Abnormal Returns because it is a direct measure of investor attention and provides real-time data instead of capturing shifts in investor attention with the delays associated with corporate data releases. Therefore, using ASVI can help investors detect shifts in attention before they are incorporated into the market.

Hypothesis II: The Abnormal Google Search Volume Index predicts Abnormal Returns in the short term.

Hypothesis III: The Abnormal Google Search Volume Index has a stronger effect on Abnormal Returns in technology stocks relative to non-technology stocks.

These last two hypotheses test whether ASVI predicts Abnormal Returns in the short term. Hypothesis II focuses on studying the predictability of Abnormal Returns based on ASVI in the short term, and Hypothesis III does the same but analyses this effect by comparing technology and non-technology stocks. I test these hypotheses on a dataset that contains all the stocks within the Nasdaq-100 and that goes from 2019 to 2023. This time window includes three periods: before, during and post-COVID-19 pandemic. The pre-COVID-19 period consists of relative market stability, the period during COVID-19 is marked by uncertainty and volatility and the post-COVID-19 period is when the market started to recover but also when other factors started to impact the market, namely the widespread adoption of remote work, wars and fluctuating government policies.

Therefore, this thesis expands the market efficiency research by studying the Abnormal Google Search Volume Index as a direct and accessible measure of investor attention, which demonstrates that ASVI provides unique insights as a proxy for investor attention, allowing retail investors to make more thoughtful choices in the stock market.

Second, this study demonstrates that ASVI can predict short-term Abnormal Returns in the Nasdaq-100 index since their relationship is positive and significant across several models when looking at the full period.

Third, this paper extends the literature by focusing on the Nasdaq-100, a US index that contains some of the most traded and well-known stocks, which normally already attract more attention (Barber & Odean, 2008). To further explore the implications of stocks that have higher levels of attention, the paper also distinguishes between technology and non-technology stocks, since the Nasdaq-100 has a good representation of both, and it is find that for technology stocks the relationship between ASVI and Abnormal Return is stronger compared to non-technology stocks, which can be explained by the higher volatility and growth compared to other sectors (Cutter, 2024; Moran, 2020), that consequently lead to more attention according to Padungsaksawasdi et al. (2019), which leads to higher Abnormal Returns (Da et al., 2011).

Building upon the main findings, this thesis further studies the effects of different market conditions. Focusing on each specific period between 2019 and 2023, it studies periods of stability, uncertainty, and recovery related to the COVID-19 pandemic, which, to my knowledge, has not been done before. The results show that during and after the COVID-19 pandemic, ASVI demonstrates its strongest predictive power for Abnormal Return, enhancing the understanding of investor behaviour during periods of economic instability. This result reflects the higher volatility and uncertainty during the pandemic (Basuony et al., 2021; Kayani et al., 2024; Ullah et al., 2023). Therefore, this shows that attention-driven effects intensify during periods of high volatility and uncertainty, aligning with studies like Bondt & Thaler (1985), which demonstrate that the market overreacts to unexpected events. Lastly, I conclude that excluding 'noisy' firms with low Google search volume hardly changes the results in Hypothesis III, confirming ASVI's stronger effect on the Abnormal Returns of technology stock.

The rest of this thesis is organized as follows. Section II describes the data. Section III explains the methodology, Section IV presents the empirical results. Section V discusses the results, and Section VI summarises and concludes this thesis.

II. Data

Variables

Following Da et al. (2011), I am using the dependent variable, Abnormal Return, and as main independent variable, ASVI. Moreover, I am using Log Abnormal Turnover, Log Size, Analyst Coverage, Advertising/Sales, Log IVOL, Idiosyncratic Skewness, News Coverage, Earnings Dummy and Tech Dummy as the control variables.

Table I
Variable Definitions

Variable	Definition	Description and Source
Abnormal Return	Abnormal Return	The difference between actual returns and the expected return derived from the Fama and French three-factor model | CRSP and Kenneth French Data Library
ASVI	Abnormal Search Volume Index	The natural logarithm of SVI during the week minus the natural logarithm of the median SVI for the previous 8 weeks | Google Trends
SVI	Search Volume Index	Scaled measure of weekly aggregate search frequency based on the following keyword: company name + stock | Google Trends
Log Abnormal Turnover	Log Abnormal Turnover	The natural logarithm of the current week's stock turnover minus the natural logarithm of the median stock turnover for the previous 8 weeks | CRSP
Log Size	Log Size	The natural logarithm of price multiplied by shares outstanding | CRSP
Analyst Coverage	Analyst Coverage	The number of analysts following each company as measured by the latest number of EPS forecasts | IBES
Advertising/Sales	Advertising/Sales	The natural logarithm of advertising expenses over total sales based on the previous fiscal year, if missing, the advertising expenses where set to zero | Compustat
Log IVOL	Log Idiosyncratic Volatility	The natural logarithm of the standard deviation of residuals after regressing weekly excess returns on the Fama French 3 factor model | CRSP and Kenneth French Data Library
Idiosyncratic Skewness	Idiosyncratic Skewness	The skewness of residuals after regressing weekly excess returns on the Fama French 3 factor model | CRSP and Kenneth French Data Library
News Coverage	News Count	The total number of news articles published about each stock per week | Refinitiv
Earnings Dummy	Earnings Announcement	Dummy variable that equals to one when earnings are announced and equal to zero otherwise | Compustat
Tech Dummy	Tech Sector	Dummy variable that equals to one when the stock belongs to the tech sector equal to zero otherwise | Compustat

The data is based on all constituents of the Nasdaq-100 between 2019 and 2023, even if the companies only joined the index later or went out of the index before, to prevent survivorship bias. Therefore, the sample has 143 stocks that at least at some point belonged to the Nasdaq-100, an index that includes the 100 biggest American non-financial stocks traded on the Nasdaq-100 stock market. This index includes some of the most well-known companies, such as Microsoft, Apple, Amazon, Nvidia, and Tesla. Besides these, there are many more companies, from very different sectors, with the technology sector representing 50% of the index.

I focus on the period from 2019 to 2023 since the study emphasises the COVID-19 period. The access to weekly data from Google Trends is restricted to a maximum period of five years, which cannot be overcome by overlapping different series because SVI is a relative measure. I choose the weekly frequency because it offers a balanced view of short-term patterns without the excessive noise, which aligns well with the literature.

I retrieve the data from several databases, such as CRSP, Compustat, IBES, Kenneth French Data Library, Refinitiv, and Google Trends as summarised in Table I. I choose the Google Trends platform because it is based on data from Google, which is the dominant platform with a market share fluctuating between 91.38% and 93.37% from 2019 to 2023 (Bianchi, 2024). The specifications used are searches in the entire world since an investor does not need to be in the US to invest in companies from the Nasdaq-100. Furthermore, I use the default specifications, including all categories and web searches.

Unlike Da et al. (2011), I use the company names combined with the word 'stock' as keywords (e.g. Microsoft stock) because it is more common for a retail investor to know the name of the company and just add the word stock compared to knowing the respective ticker, as confirmed in Google Trends. Additionally, the name of the company combined with the word 'stock' reflects financial searches rather than people trying to reach the company website or trying to buy products/services, depending on the company. I exclude words such as 'Inc', 'Corp', 'Co', 'Ltd', '.Com', 'Plc', 'Nv', and 'Holding' from the company names because frequently retail investors do not include them when searching for each company, as confirmed in Google Trends. I also account for exceptional situations where a company changes its name, and for that, I retrieve the values for the different names in the same search to be able to sum them since the SVI is a relative measure. These and other cases where I make minor changes to align with the investor's research can be identified in Appendix I.

Abnormal Return

I define Abnormal Return as a return that cannot be explained by broad market factors (market, size, and value) (Fama & French, 1993). This variable provides both the magnitude and direction of deviation from expected returns. I prefer this approach to the Characteristic-adjusted returns approach used by Da et al. (2011) due to the unique characteristics of the period being studied since the Fama-French abnormal returns approach accounts for macroeconomic shocks, which were prominent in the dataset.

$$\text{Abnormal Return} = \text{Actual Return} - \text{Expected Return}$$

$$\text{Expected Return} = \alpha + \beta_1 (\text{Mkt-Rf}) + \beta_2 (\text{SMB}) + \beta_3 (\text{HML})$$

Abnormal Google Search Volume Index (ASVI)

Following Da et al. (2011), I use the ASVI as the main independent variable since it measures the unusually high levels of searches compared with each typical value. To ensure that ASVI reflects truly the deviation from the typical values, I use the median of the last 8 weeks, which provides a stable benchmark for typical values of attention while smoothing out short-term fluctuations. Lastly, I apply the log transformation to reduce the influence of outliers and right skewness in the data. Since SVI is a relative measure between 0 and 100, and some values are zero, which would cause undefined log transformations, I also add one inside each logarithm formula. Besides the ASVI that is studied by Da et al. (2011), the SVI has also proven its value in the literature several times by showing that investor attention is strongly correlated with Trading Volume, Turnover, Stock Liquidity and Volatility (Aouadi et al., 2013; Ayala et al., 2024; Bank et al., 2011).

$$ASVI_t = \ln(1 + SVI_t) - \ln[1 + Med(SVI_{t-1}, \ldots, SVI_{t-8})]$$

Besides the dependent variable, Abnormal Return, and the main independent variable, ASVI, I also include the following variables to minimise the omitted variable bias.

Log Abnormal Turnover

The variable represents a transformed measure of unusual trading volume by the number of shares outstanding relative to its recent values, meaning it captures deviations in trading activity from a stock's usual levels. This variable is crucial as a proxy for investor attention because while ASVI reflects shifts in investor interest online, Log Abnormal Turnover focuses on the trading behaviour in the markets, so by including this variable it is possible to determine if the relationship between ASVI and Abnormal Returns is truly driven by attention itself or if the effect is due to higher trading volume. To compute the Log Abnormal Turnover, I use the median over the past 8 weeks instead of the mean in order to reduce the impact of extreme values and to ensure the data is sensitive to recent changes. Once again, I apply a log transformation to provide a proportional measure of Abnormal Turnover, which helps in capturing relative changes. I use Log Abnormal Turnover instead of Standardized Abnormal Turnover as in Da et al. (2011) to focus on the short term and to reduce skewness and smooth outliers, which are critical given the period being studied.

$$\text{Turnover} = \text{Trading Volume}/\text{Shares Outstanding}$$

$$\text{Log Abnormal Turnover} = \ln(\text{Turnover}_{it}) - \ln[Med(\text{Turnover}_{it-1}, \ldots, \text{Turnover}_{it-8})]$$

Log Size

As in Da et al. (2011), I use the variable Market Capitalisation, but I apply the log transformation to reduce skewness and provide an interpretable measure of firm size.

$$\text{Market Capitalisation} = \text{Price} * \text{Shares Outstanding}$$

$$\text{Log Size} = \ln(\text{Market Capitalisation})$$

Analyst Coverage

As in Da et al. (2011), the variable represents the number of financial analysts who follow and report about a particular company. Since Analyst Coverage is only available monthly in IBES and the monthly changes are small, I assign the weekly values equal to the monthly values of the respective month.

Advertising/Sales

As in Da et al. (2011), this variable measures the company's advertising expenses relative to its sales, but I apply the log transformation to reduce skewness. It is important to control for this variable because according to Grullon et al. (2004), advertising serves to communicate new products but also to increase the firm's visibility among investors, which can lead to higher Abnormal Returns. However, this variable has some limitations. To begin with, there are a lot of missing values regarding advertising expenses, so I assume that missing values are zero to maintain sample size and simplify the analysis, as in previous studies (e.g., Da et al., 2011; Ding & Hou, 2015). This assumption is also supported by Grullon et al. (2004) and Lou (2014) who state that assuming the missing values as zero is qualitatively similar to only select firms that report advertising expenses. However, this leads to some bias because it is not possible to distinguish between firms that have zero advertising and firms that do not want to disclose their values. Additionally, the data about Advertising Expenses and Sales retrieved from Compustat is only available annually, so I fill in the annual value for each week of the respective year. This can also introduce bias because the Advertising Expenses and Sales can change within a year which leads to less precise results. Finally, since many Advertising Expenses are zero, which would lead to errors in the logarithm formula, I also add one to the ratio of advertising expenses over sales.

$$\text{Advertising/Sales} = \ln(1 + \text{Advertising Expenses/Sales})$$

Log Idiosyncratic Volatility (IVOL)

Since Da et al. (2011) do not use any volatility variable as a control variable and following Kumar (2009), I use this variable because high idiosyncratic volatility often attracts speculative investors seeking large prospective gains. This variable represents firm-specific volatility that is not explained by market factors and is used because it enables the understanding of the stock's distinctive risk characteristics and its relationship to investor attention, without considering the influence of broader market dynamics. To capture quick shifts while smoothing the noise caused by isolated events, I apply the natural logarithm, which accounts for extreme values.

$$\text{Log IVOL} = \ln(\text{StDev.S}(\text{Abnormal Return}_{t-8}, \ldots, \text{Abnormal Return}_{t-1})$$

Idiosyncratic Skewness

Since Da et al. (2011) do not use any skewness variable as a control variable, and following Kumar (2009), this variable measures the asymmetry of a stock's return distribution, capturing the likelihood of extreme positive or negative events, which often influence investor behaviour and market dynamics.

$$\text{Idiosyncratic Skewness} = \text{SKEW}(\text{Abnormal Return}_{t-8}, \ldots, \text{Abnormal Return}_{t-1})$$

News Coverage

This variable represents the amount of news as in Da et al. (2011) but from a different source since their source is paid. Contrary to ASVI, this variable reflects the quantity of published news rather than investor attention. I apply the log transformation to reduce skewness and provide an interpretable measure of firm size. But since some values are zero, which would cause undefined log transformations, I also add one inside the logarithm. To retrieve the data, ideally, I would use the company names instead of the tickers as keywords because I am not only interested in news posted in financial newspapers, but since free data is unavailable, I retrieve the number of news articles from Refinitiv using the tickers as keywords, which underestimate the overall quantity of published news receive. Additionally, Refinitiv does not maintain historical archives for delisted companies. So, I cannot retrieve data for delisted companies from the Nasdaq-100 due to either a merger or acquisition. These can be a limitation if these companies have a particular pattern of investor attention.

$$\text{Log News Coverage} = \ln(1 + \text{News Coverage})$$

Since bigger firms often attract more Analyst and News Coverage, I test for multicollinearity between Log Size and Analyst Coverage, and between Log Size and News Coverage. To do that I use the Variance Inflation Factor (VIF) and examine the correlation matrix in Table III. The correlation between Log Size and Analyst Coverage is 0.400 and the VIF is 1.03. The correlation between Log Size and News Coverage is 0.431 and the VIF is 1.04. These values mean that each variable provides relatively independent information and therefore their inclusion does not compromise the regressions' reliability.

Earnings Announcement Dummy

Differently from Da et al. (2011), I include the Earnings Announcement Dummy because earnings announcements are events that often lead to an increase in investor attention and substantial price volatility, and therefore can be the reason for such returns around those periods. Drake et al. (2012) prove this in their paper where they found that ASVI increases before the earnings announcement, spikes at the announcement, and continues high during a certain period. However, they also found that when investors are paying more attention to the earnings announcement the information is partially absorbed in advance.

Tech Dummy

I also include a Tech Dummy to test Hypothesis III and examine whether technology stocks behave differently from non-technology stocks, differently from Da et al. (2011). Since technology stocks often have more innovation cycles, growth potential and high volatility, which attracts the attention of more investors (Cutter, 2024; Moran, 2020).

Other Variables

Besides the variables described above, and as in Da et al. (2011), I also include lagged variables, namely Lagged Abnormal Return, Lagged ASVI, Lagged Abnormal Turnover and Lagged News Coverage, to see how past events influence current outcomes and therefore address several issues such as autocorrelation, endogeneity, and reverse causality.

All variables used in this thesis are described in Table I, and the respective descriptive statistics are provided in Table II. Additionally, Table III and Appendix II, III, and IV show different correlation matrices between all the above variables in the different periods.

Descriptive Statistics

Table II presents the descriptive statistics for the entire sample period and each of the three sub-periods. The variable SVI has a higher mean and median during and post-COVID-19, probably because investors were more concerned with the economic outlook and had more time to look to the stock market to make better investment decisions (Yahya et al., 2021). The Log Abnormal Turnover is also higher during and post-COVID-19, which means the trading activity increased, possibly due to the volatile conditions of those periods.

Moreover, I find that almost all variables are moderately skewed, either positively or negatively, with most being positive. ASVI and Advertising/Sales have positive skewness, meaning that the majority of the data are small values, and few of them are much larger, pulling the mean towards the right of the distribution. On the contrary, the variable Abnormal Return is negatively skewed during most of the period, indicating that extreme negative returns are more frequent than extreme positive ones, which pulls the mean towards the left of the distribution. These extreme outliers affect the regressions' results which assume a symmetric distribution.

As seen in Table II, the limitations mentioned above regarding the variable Advertising/Sales have a visible impact since by assuming missing values as zero, the variable has positive skewness, and the mean and median are really close to zero, which means that the variable may have limited explanatory value.

Lastly, all the variables present excess kurtosis higher than zero which indicates that the data's distribution has very fat tails suggesting the presence of extreme values, which was expected since the majority of the period that is studied is marked by a period of uncertainty due to the expectations related to the virus' progression. These extreme values can reduce the model's precision if the models assume normality as happens in the case of Pooled OLS and FE Panel Data. The models CCEMG and GMM, on the contrary, already mitigate the effect of fat tails but do not eliminate them.

Table II
Descriptive Statistics

Variable	Period	N	Mean	Median	Std. Dev.	Min	Max	Skewness	Excess Kurtosis
Abnormal Return	01/2019 - 01/2020	5663	0.037	0.037	0.042	-0.469	0.467	-0.067	14.781
	02/2020 - 04/2020	2261	0.041	0.038	0.054	-0.289	0.482	1.029	9.994
	05/2020 - 12/2023	26064	0.037	0.036	0.052	-0.201	0.672	-5.518	16.549
	Full	33988	0.037	0.037	0.051	-0.201	0.672	-4.536	14.077
Search Volume Index	01/2019 - 01/2020	5663	17.485	15.000	15.756	0.000	100.000	1.060	1.704
	02/2020 - 04/2020	2261	32.636	29.000	24.160	0.000	100.000	0.729	0.075
	05/2020 - 12/2023	26064	22.494	20.000	18.363	0.000	100.000	1.002	1.286
	Full	33988	22.334	19.000	18.702	0.000	100.000	1.056	1.422
Abnormal Search Volume Index	01/2019 - 01/2020	5663	0.026	0.000	0.845	-4.174	4.615	0.485	10.776
	02/2020 - 04/2020	2261	0.280	0.102	0.836	-4.263	4.615	1.752	10.974
	05/2020 - 12/2023	26064	0.005	0.000	0.762	-4.263	4.615	0.288	14.848
	Full	33988	0.269	0.000	0.785	-4.263	4.615	0.458	13.686
Log Abnormal Turnover	01/2019 - 01/2020	5663	-0.004	-0.029	0.376	-1.486	2.453	0.639	2.339
	02/2020 - 04/2020	2261	0.130	0.087	0.472	-1.592	2.536	0.341	0.306
	05/2020 - 12/2023	26064	0.004	-0.022	0.366	-1.533	2.386	0.623	2.024
	Full	33988	0.011	-0.018	0.377	-1.592	2.536	0.628	1.917
Log Size	01/2019 - 01/2020	5663	17.260	17.098	1.100	13.678	20.963	0.632	1.489
	02/2020 - 04/2020	2261	17.263	17.045	1.135	13.670	21.075	0.720	1.322
	05/2020 - 12/2023	26064	17.652	17.485	1.096	14.202	21.846	1.090	1.925
	Full	33988	17.561	17.406	1.111	13.670	21.846	0.945	1.828
Analyst Coverage	01/2019 - 01/2020	5663	21.108	21.000	8.903	0.000	48.000	-0.113	0.233
	02/2020 - 04/2020	2261	20.651	21.000	9.817	0.000	46.000	-0.208	-0.048
	05/2020 - 12/2023	26064	19.953	21.000	10.036	0.000	48.000	-0.249	-0.137
	Full	33988	20.192	21.000	9.851	0.000	48.000	-0.240	-0.060
Advertising/Sales	01/2019 - 01/2020	5663	0.028	0.004	0.067	0.000	0.620	5.900	45.401
	02/2020 - 04/2020	2261	0.027	0.004	0.056	0.000	0.507	5.287	38.497
	05/2020 - 12/2023	26064	0.028	0.005	0.052	0.000	0.507	3.662	19.139
	Full	33988	0.028	0.004	0.055	0.000	0.620	4.493	30.851
Log Idiosyncratic Volatility	01/2019 - 01/2020	5663	-3.488	-3.509	0.567	-5.247	-1.526	0.202	0.036
	02/2020 - 04/2020	2261	-3.424	-3.412	0.601	-5.322	-1.493	0.147	-0.016
	05/2020 - 12/2023	26064	-3.363	-3.390	0.557	-5.455	-0.075	0.444	0.981
	Full	33988	-3.388	-3.411	0.564	-5.455	-0.075	0.369	0.759
Idiosyncratic Skewness	01/2019 - 01/2020	5663	0.007	0.004	1.015	-2.777	2.744	-0.083	-0.116
	02/2020 - 04/2020	2261	0.087	0.102	0.940	-2.739	2.587	-0.206	0.224
	05/2020 - 12/2023	26064	0.061	0.069	0.957	-2.823	2.791	-0.098	0.090
	Full	33988	0.054	0.062	0.966	-2.823	2.791	-0.105	0.062
News Coverage	01/2019 - 01/2020	5662	1.727	1.609	1.583	0.000	6.938	0.664	-0.301
	02/2020 - 04/2020	2261	1.861	1.792	1.552	0.000	7.241	0.581	-0.271
	05/2020 - 12/2023	26063	1.990	1.792	1.677	0.000	7.854	0.635	-0.264
	Full	33986	1.938	1.792	1.656	0.000	7.854	0.642	-0.252

Correlation Matrix

Table III, along with Appendix II, III, and IV, present the correlation matrices for all variables across the different periods. The Table III for the full period, the Appendix II from 01/2019 to 01/2020, the Appendix III from 02/2020 to 04/2020, and the Appendix IV from 05/2020 to 12/2023. The aim is to capture how the different variables correlate with each other and with the market over time.

The Log Size seems to be the variable with higher correlations with other variables. It is moderately negatively correlated with Log IVOL, -25.4%, which is expected since more stable firms tend to have lower idiosyncratic volatility, and it is moderately positively correlated with the variable Analyst Coverage, even more during the COVID-19 period, 49.4%, possibly due to increased scrutiny of larger firms during periods of major instability. Additionally, it is moderately positively correlated with the variable News Coverage, 43.1%, which means that as expected, larger firms have more news coverage due to its recognition. However, as shown before there are no reasons to exclude one of them since there are no multicollinearity issues.

Furthermore, the low coefficient between ASVI and News Coverage, 5.1%, supports the idea that News Coverage is not directly correlated with investor attention. This can be attributed to the fact that only 42% of Nasdaq-100 stocks receive coverage in a given year (Fang & Peress, 2009), and an increase in the number of news does not necessarily mean they are being read (Da et al., 2011). However, this result also reflects the limitations in the Refinitiv dataset, which underrepresent broader news coverage.

In conclusion, I find that the lower correlations between ASVI and other proxies of investor attention suggest that ASVI capture unique aspects of investor attention not explained by other variables, as will be then confirmed in Hypothesis I.

Table III
Correlation Matrix

	Abnormal Return	ASVI	Log Abnormal Turnover	Log Size	Analyst Coverage	Advertising/Sales	Log IVOL	Idiosyncratic Skewness	News Coverage
Abnormal Return	1.000								
ASVI	0.027	1.000							
Log Abnormal Turnover	0.040	0.108	1.000						
Log Size	0.017	-0.006	-0.022	1.000					
Analyst Coverage	-0.020	-0.005	-0.007	0.400	1.000				
Advertising/Sales	0.008	-0.005	0.006	-0.124	-0.057	1.000			
Log IVOL	0.007	-0.057	-0.046	-0.254	-0.154	0.174	1.000		
Idiosyncratic Skewness	0.007	-0.007	-0.028	0.001	-0.027	0.010	0.022	1.000	
News Coverage	0.030	0.051	0.198	0.431	0.258	0.018	0.021	0.009	1.000

III. Methodology

Variance Inflation Factor

Unlike Da et al. (2011), I run a Variance Inflation Factor (VIF) in Stata to ensure that there are no multicollinearity issues affecting the reliability of the results. This test checks if any independent variables are highly correlated, which could compromise the regression's validity. The results show that all variables have VIF below 5, so according to Akinwande et al. (2015), there is no indication of a high correlation that indicates the presence of multicollinearity. The largest VIF is equal to 1.52, so the VIF values are low, and even though moderate correlations may still introduce some overlap in explaining variance, they do not invalidate the model, and therefore there is no need to drop any variables from the model.

Fixed vs Random Effects

Thereafter, I use the Hausman test in Stata to decide between the Fixed Effects (FE) model and the Random Effects (RE) model. The Hausman test determines if the firm and year-specific characteristics are correlated with the independent variables. If they are correlated, the RE model cannot be used due to the risk of omitted variable bias since the model would fail to control for these unobserved factors adequately, making the FE model the most appropriate choice. (Hausman, 1978)

In this thesis, the Chi-squared associated with this test is negative which means the model does not meet the assumptions needed for the Hausman test. Therefore, it is necessary to perform the adjusted Hausman test (sigamore) that corrects the scaling differences. In this adjusted test, the Chi-squared value is 373.59 and I can reject the null hypothesis at a 1% significance level, which means I reject the statement that there are no correlations between individual effects and explanatory variables. Therefore, the FE model is the most adequate, as in Da et al. (2011), because it adequately controls for the correlation detected by the Hausman test. This thesis uses firm FE to focus on within-firm variations over time, controlling for unobserved time-invariant factors like management style and market positioning. This allows the study of the independent variables' impact on Abnormal Return without interference from firm-specific characteristics. Additionally, this thesis uses year-fixed effects to account for time-varying factors common across firms, such as COVID-19, monetary policy changes, and economic recessions.

To sum up, the FE model controls for time-invariant firm-specific characteristics and common time-varying shocks, but it does not account for unobserved time-varying shocks specific to individual firms, which can influence the results. (Hausman, 1978)

Pesaran's CD Test

Considering the results from the Hausman Test and the period studied in this paper, I analyse whether there is cross-sectional dependence, diverging from Da et al. (2011). Cross-sectional dependence arises from unobserved common shocks, and this occurs when the residuals from different companies are correlated due to events that affect all companies simultaneously (Bernard, 1987). However, since the panel data is unbalanced, I use Pesaran's CD Test for unbalanced panels. An unbalanced panel is a set of data that does not begin and ends at the same dates and/or has missing values. This situation applies to the dataset since I study all firms that belonged at some point to the Nasdaq-100 between 2019 and 2023, and some of them only entered the index after 2019 or left before 2023. So, therefore, the unbalanced panel is necessary in order to overcome the survivorship bias problem.

The result from the Pesaran's CD Test is 27.09, and I can strongly reject the null hypothesis of weak cross-sectional dependence at a 1% significance level. This result indicates that the residuals from different companies are correlated, which suggests that there are common factors that affect all the companies at the same time, in other words, I find strong cross-sectional dependence in the residuals, possibly due to the COVID-19 pandemic or the monetary policy shifts.

Since there is strong cross-sectional dependence, methods such as Fama-MacBeth used by Da et al. (2011) that assume independence of errors across companies should not be used, so I use different methods to test Hypothesis II and III (Gow et al., 2010). Additionally, for Hypothesis I, the Vector Autoregressions (VAR) method used by Da et al. (2011) is not used because it is best suited for time-series data, assumes stationarity over time, and lacks robustness compared to other models due to the particular characteristics of this dataset. Therefore, the methods that best suit the dataset are the Pooled OLS, the FE Panel Model, the Common Correlated Effects Mean Group (CCEMG), and the System Generalized Method of Moments (System GMM), which are different from the methods used by Da et al. (2011). These methods, by having different focuses, provide various perspectives on the data and enable the analysis of how the different methods' assumptions have an impact on the results.

FE Panel Model

The FE Panel Model partially addresses endogeneity by controlling for unobserved time-invariant factors, such as management style and market positioning. Additionally, I include year-fixed effects to account for time-varying factors common across firms even though the model does not fully account for this as it cannot capture unobserved time-varying factors unique to individual firms. Besides the FE Panel Model, the only model that uses fixed effects is the Pooled OLS which includes firm and year-fixed effects. To test the reliability of these methods, I compute the R-squared which measures the proportion of the variation in the dependent variable explained by the independent variables. The R-squared is well suited for these methods because they aim to minimise residual variance, which the R-squared directly reflects. Nevertheless, since the FE Panel Model and the Pooled OLS assume normality and the dataset has high excess kurtosis, other models, such as CCEMG and GMM, that do not assume normality are essential because they are best suited to this dataset.

Common Correlated Effects Mean Group

The CCE estimator is divided into the CCE Mean Group (CCEMG) and the CCE Pooled (CCEP) estimators. Since I detect cross-sectional dependence, I choose to use only the CCEMG because it performs better when there are unobserved common shocks and heterogeneous responses across firms (Ditzen, 2018; Pesaran, 2006).

The CCEMG addresses partially endogeneity by capturing the unobserved common shocks but does not account for reverse causality or autocorrelation. Unlike the FE model, the CCEMG does not explicitly control for time-invariant factors but indirectly accounts for them by focusing on cross-sectional averages, which capture the effect of common shocks. Lastly, I use the CCEMG to indirectly address time-varying factors and heteroskedasticity by capturing how firms respond heterogeneously to common shocks.

Since the CCEMG's objective is to correct for cross-sectional dependence and not to minimise the residual variance, the R-squared is not meaningful. Therefore, to ensure robustness I use the Root Mean Squared Error (RMSE) as it measures the prediction error and is more aligned with this model.

System Generalized Method of Moments

Lastly, I use the System Generalized Method of Moments (GMM) to capture endogeneity with dynamic effects. Since the dataset probably reflects the influence of previous observations, I find it essential to include lagged values.

First, the GMM model addresses endogeneity, caused by omitted variables and reverse causality, by using lagged values of endogenous variables. In addition, it also addresses autocorrelation caused by lagged dependent variables and reverse causality by identifying causal relationships when the dependent and independent variables influence each other. (Arellano & Bover, 1995)

Second, it partially addresses time-invariant factors, and accounts for time-varying shocks which are especially relevant in volatile markets like the Nasdaq-100 since the fluctuations lead to frequent and unpredictable, shocks that can impact stock returns (Arellano, 1991).

Finally, GMM's robustness to heteroskedasticity allows it to handle situations where the variance of the error term is not constant across observations, which is likely in this dataset (Arellano, 1991).

Since the System GMM aims to address endogeneity and not to minimise the residual variance, the R-squared is also not meaningful. Therefore, to ensure robustness I use the Arellano-Bond test and Hansen/Sargan test for GMM validity. These tests address specific issues relevant to GMM models, such as correlation and instrument validity.

IV. Empirical Results

This section explains each hypothesis defined in the Introduction and presents the respective empirical results.

Hypothesis I: ASVI and Investor Attention Proxies

I study this hypothesis first because ASVI is not very useful if it can already be explained by other investor attention proxies, such as Trading Volume (Barber & Odean, 2008; Gervais et al., 2001; Preis et al., 2010), Advertising/Sales (Grullon et al., 2004; Lou, 2014), Analyst Coverage (Da et al., 2011), and News Coverage from journals (Barber & Odean, 2008; Tetlock, 2007; Thompson II et al., 1987). In this analysis, I test whether the ASVI serves as a unique proxy for investor attention compared to the remaining investor attention proxies, since ASVI has the advantage of being based on search behaviour, and then it can capture retail investors' attention directly and quicker.

As in Da et al. (2011), I test whether ASVI is identical to other investor attention proxies, but I use different econometric models: Pooled OLS, FE Panel Model, CCEMG, and GMM. I use the Pooled OLS and FE Panel Model as a baseline since they do not address cross-sectional dependence or endogeneity. To address these issues, I use the CCEMG of Pesaran (2006) to handle cross-sectional dependence, and the GMM to address endogeneity and time-varying shocks (Arellano & Bover, 1995). As explained in the Methodology, I avoid the Vector Autoregressions (VAR) method used by Da et al. (2011), because it is best suited for time-series data, assumes stationarity over time, and lacks robustness compared to other models.

Table IV
ASVI and Investor Attention Proxies

	(1) Pooled OLS	(2) FE Panel Model	(3) CCEMG	(4) System GMM
Abs_Abnormal_Return	0.900^{***}	0.879^{***}	1.042^{***}	0.825^{***}
	(0.103)	(0.103)	(0.104)	(0.111)
Log_Abnormal_Turnover	0.174^{***}	0.168^{***}	0.136^{***}	0.087^{***}
	(0.012)	(0.012)	(0.019)	(0.026)
Analyst_Coverage	-0.001^{**}	0.002	0.003	-0.001
	(0.001)	(0.002)	(0.004)	(0.001)
Log_Size	-0.023^{***}	-0.017	-0.015	-0.020^{**}
	(0.005)	(0.012)	(0.063)	(0.008)
AdvertisingSales	-0.006	-0.130	0.011	0.075
	(0.078)	(0.239)	(6.648)	(0.124)
Log_IVOL	-0.100^{***}	-0.118^{***}	-0.109^{***}	-0.094^{***}
	(0.008)	(0.010)	(0.016)	(0.014)
Idiosyncratic_Skewness	-0.005	-0.005	-0.001	-0.006
	(0.004)	(0.004)	(0.007)	(0.005)
News_Coverage	0.023^{***}	0.027^{***}	0.016^{***}	0.032^{***}
	(0.003)	(0.004)	(0.005)	(0.006)
Earnings_Dummy	-0.002	-0.004	-0.152^{**}	-0.008
	(0.017)	(0.017)	(0.065)	(0.014)
Lagged_ASVI				0.105^{***}
				(0.031)
Lagged_Abnormal_Return				0.236^{**}
				(0.118)
Lagged_Log_Abnormal_Turnover				0.070^{***}
				(0.024)
Lagged_News_Coverage				-0.018^{***}
				(0.007)
_cons	0.015	-0.217	1.566	-0.000
	(0.077)	(0.214)	(1.046)	(0.137)
Observations	33985	33985	33985	33839

Standard errors in parentheses, * $p < 10\%$, ** $p < 5\%$, *** $p < 1\%$

In this hypothesis, I use Abs Abnormal Return instead of Abnormal Return because the objective is to capture the magnitude of the market movements, regardless of the direction. Besides that, using absolute values prevents positive and negative Abnormal Returns from cancelling each other, meaning that all unexpected changes that attract investor attention are being considered. Across all models, Abs Abnormal Return, Log Abnormal Turnover, Log IVOL, and News Coverage consistently reveal significant effects on ASVI. The significant but negative coefficient of Log IVOL may indicate that investors tend to avoid stocks perceived as risky or difficult to evaluate. Conversely, the significant and positive coefficients suggest that investors are more prone to search for stocks with more news, that have increases in trading activity, and have recent extreme, either positive or negative, market performance, since this can lead investors to search for more information. The variable Analyst Coverage, for instance, exhibits limited significance, likely because the firm-fixed effects absorb much of the variation, as the model focuses on within-changes.

The R-squared from the Pooled OLS means that only 2.05% of the variation in the ASVI is explained by the other investor attention proxies, and the Within R-squared from the FE Panel Model means that only 2.11% of the within-group variation in the ASVI is explained by other investor attention proxies since the between-group variation is absorbed by the fixed effects. Therefore, the R-squared from both methods are not comparable but both are very low which indicates that the other investor attention proxies only explain a small portion of the variation in ASVI. However, it cannot be excluded the possibility that these low values happen due to omitted variables. The RMSE value from CCEMG of 0.7039 reflects moderate prediction accuracy when evaluated against the distribution of ASVI values in the dataset, which range from -4.26 to 4.62. This indicates that the model captures most of the variation in ASVI, though some error persists. To support the validity of the GMM model, the following tests were performed: The Arellano-Bond Test for Serial Correlation and the Sargan and Hansen Test for Overidentification. The AR(1) test is significant at a 1% significance level, confirming first-order serial correlation, but the AR(2) test is also significant at a 5% significance level, which is not desirable, as it suggests a violation of the GMM assumption that instruments should not be correlated with the error term. However, Hansen's p-values, 0.247 and 0.569, ensure that the instruments used in the model are likely uncorrelated with the error term. These contradictions are justified by the exceptional market conditions during COVID-19, such as the increased use of digital platforms (Statista Research Department, 2023), volatile market conditions and uncertainty (Basuony et al., 2021; Kayani et al., 2024; Ullah et al., 2023), which contribute to short-term data distortions. Consequently, the GMM results have to be interpreted with caution, because the patterns in ASVI are not fully captured by the control variables, which makes the CCEMG the most suited method.

To sum up, given the consistent significance of certain variables, and the low explanatory power of the other investor attention proxies, I reject the null hypothesis, as in Da et al. (2011), meaning that ASVI, as a proxy for investor attention, captures unique insights into investor attention that other proxies cannot explain, at least in the Nasdaq-100 and during the period studied, so the same results may not generalise to indices with different compositions and periods with different conditions.

Hypothesis II: The Predictive Power of ASVI on Abnormal Return

In this section, I test whether the ASVI, as a proxy for investor attention, has a statistically significant impact on Abnormal Return in the Nasdaq-100. However, I focus on the short-term effects, and I do not study whether the price reverses in the long term, differing from Da et al. (2011). Regarding the short-term results, I expect a positive impact of ASVI on Abnormal Returns since retail investors normally are uninformed and are buyers of attention-grabbing stocks which leads to temporarily higher returns (Barber & Odean, 2008).

I built on the work of Da et al. (2011) by using four regression models, instead of the Fama-MacBeth cross-sectional regression (Fama & MacBeth, 1973) and the Newey-West adjustment with eight lags (Newey & West, 1987), because Fama and Macbeth (1973) assume independence of errors across companies, focus on cross-sectional variation across firms at each point in time and assume stable relationships. Since this study focuses on the COVID-19 period which is characterised by strong cross-sectional dependence and varying market volatility, it is likely that the relationship between ASVI and Abnormal Return changes across periods, so the Fama-Macbeth is not the suited method (Gow et al., 2010).

Therefore, I use the Pooled OLS, which provides a baseline for understanding the relationship between ASVI and Abnormal Return. The FE Panel Model, that controls for time-invariant characteristics of each firm and isolates the within-firm variation over time (Baltagi, 2008). The CCEMG, which addresses cross-sectional dependence, accounting for possible correlations between firms due to unobservable or market-wide shocks (Pesaran, 2006), and the System GMM, that controls for endogeneity through the use of lagged values (Arellano & Bover, 1995).

Table V
Effect of ASVI on Abnormal Return

	(1) Pooled OLS	(2) FE Panel Model	(3) CCEMG	(4) System GMM
ASVI	0.145***	0.154***	0.703***	0.105***
	(0.040)	(0.040)	(0.120)	(0.030)
Log_Abnormal_Turnover	0.478***	0.505***	0.320*	0.365***
	(0.080)	(0.080)	(0.160)	(0.130)
Analyst_Coverage	-0.018***	0.006	0.023	-0.017***
	(0.000)	(0.010)	(0.030)	(0.000)
Log_Size	0.132***	0.789***	6.740***	0.114***
	(0.030)	(0.080)	(0.460)	(0.040)
AdvertisingSales	0.061	1.390	208.100	-0.031
	(0.510)	(1.580)	(170.560)	(0.370)
Log_IVOL	0.093*	0.027	-0.058	0.096
	(0.050)	(0.060)	(0.140)	(0.080)
Idiosyncratic_Skewness	0.038	0.037	-0.124***	0.015
	(0.030)	(0.030)	(0.040)	(0.030)
News_Coverage	0.053***	0.055**	0.057	0.055**
	(0.020)	(0.030)	(0.050)	(0.020)
Earnings_Dummy	0.044	0.041	-0.302	0.054
	(0.110)	(0.110)	(0.460)	(0.110)
Lagged_ASVI				0.179***
				(0.060)
Lagged_Abnormal_Return				2.800
				(4.400)
_cons	2.010***	-10.000***	-17.400	2.140***
	(0.500)	(1.420)	(12.640)	(0.530)
Observations	33985	33985	33985	33842

Standard errors in parentheses, * $p < 10\%$, ** $p < 5\%$, *** $p < 1\%$

Looking at the results from the regressions in basis points (bps), because it is the standard unit to compute Abnormal Returns, I can conclude that ASVI is statistically significant across all models. Looking at each model, the Pooled OLS, the FE Panel Model and the GMM show smaller but significant coefficients, which indicate consistency but weaker relationships. The CCEMG model exhibits the strongest effect size, which confirms the immediate robustness of this relationship under cross-sectional dependence. These results support the hypothesis being tested that an increase in investor attention, captured by ASVI, predicts an increase in short-term Abnormal Returns, as studied in previous literature (Ayaz et al., 2021; Barber & Odean, 2008; Da et al., 2011; Tan & Taş, 2019). In other words, a one-standard-deviation increase in ASVI leads to a significant positive Abnormal Return change of 0.703 bps among the Nasdaq-100 stocks. This result means that stocks receiving higher attention may experience improved performance beyond their expected returns. So, investors could leverage this as part of a strategy to identify stocks that are getting more attention in the stock market, especially in high-attention environments. Nevertheless, this would have to be further studied to make sure that this leads to a profitable strategy.

Log Abnormal Turnover and Log Size have positive and significant coefficients in all models. This means that trading activity can predict Abnormal Returns (Gervais et al., 2001) as well as Log Size. The Log Size impact can be justified due to the fact that bigger firms have greater liquidity and lower perceived risk. But to confirm this capability, it would be needed to study each one of the variables as the main independent variable as in Gervais et al. (2001) who state that stocks with high trading volume for one day to a week tend to appreciate in the following month. The variable News Coverage measures the numbers of news related to each firm in Refinitiv so the lack of significance in the CCEMG model may be due to the aggregation effects, as firm-specific news impacts are not uniformly correlated across groups. Finally, the Lagged ASVI's positive and significant coefficient in the GMM model highlights that the effect of ASVI on Abnormal Returns is not limited to immediate impacts but persists over time.

Table VI
Effect of ASVI on Abnormal Return: Robustness

	(1) Pooled OLS	(2) FE Panel Model	(3) CCEMG	(4) System GMM
	Panel A. January 2019 to January 2020			
ASVI	0.195***	0.206***	0.369	0.133**
	(0.060)	(0.060)	(0.260)	(0.050)
Log_Abnormal_Turnover	-0.267*	-0.179	-0.545	-0.437*
	(0.140)	(0.150)	(0.370)	(0.240)
Analyst_Coverage	-0.016**	-0.020	0.190	-0.013*
	(0.010)	(0.020)	(0.320)	(0.010)
Log_Size	0.143**	4.750***	32.400***	0.140*
	(0.060)	(0.390)	(2.750)	(0.080)
AdvertisingSales	0.293	10.400	580.100	0.195
	(0.810)	(9.740)	(618.250)	(0.590)
Log_IVOL	0.091	-0.060	0.482	0.077
	(0.110)	(0.140)	(0.850)	(0.110)
Idiosyncratic_Skewness	-0.012	-0.143***	-0.581***	-0.017
	(0.050)	(0.050)	(0.160)	(0.050)
News_Coverage	0.092**	0.069	0.066	0.078
	(0.040)	(0.060)	(0.120)	(0.050)
Earnings_Dummy	-0.036	-0.056	-1.130	-0.068
	(0.240)	(0.240)	(0.770)	(0.200)
Lagged_ASVI				0.202**
				(0.090)
Lagged_Abnormal_Return				-3.940*
				(2.120)
_cons	1.690*	-78.500***	112.100	1.810
	(0.930)	(6.730)	(99.740)	(1.120)
Observations	6327	6327	6327	6193
	Panel B. February 2020 to April 2020			
ASVI	0.129***	0.142***	0.798***	0.086**
	(0.040)	(0.040)	(0.140)	(0.040)
Log_Abnormal_Turnover	0.650***	0.690***	0.527***	0.525***
	(0.090)	(0.090)	(0.180)	(0.150)
Analyst_Coverage	-0.018***	0.005	0.011	-0.019***
	(0.000)	(0.010)	(0.040)	(0.010)
Log_Size	0.131***	1.170***	8.300***	0.118***
	(0.040)	(0.110)	(0.560)	(0.040)
AdvertisingSales	-0.066	0.903	244.900	-0.040
	(0.620)	(2.020)	(207.350)	(0.530)
Log_IVOL	0.105*	0.053	-0.125	0.130
	(0.060)	(0.080)	(0.150)	(0.090)

Table VI - Continued

	(1) Pooled OLS	(2) FE Panel Model	(3) CCEMG	(4) System GMM
Idiosyncratic_Skewness	0.052	0.044	-0.150***	0.030
	(0.030)	(0.030)	(0.050)	(0.040)
News_Coverage	0.044*	0.058*	0.045	0.048**
	(0.020)	(0.030)	(0.060)	(0.020)
Earnings_Dummy	0.063	0.069	-0.051	0.079
	(0.120)	(0.120)	(0.590)	(0.130)
Lagged_ASVI				0.183**
				(0.070)
Lagged_Abnormal_Return				2.320
				(4.790)
_cons	2.060***	-16.600***	-21.200	2.230***
	(0.580)	(1.910)	(14.160)	(0.580)
Observations	27658	27658	27658	27516
Panel C. May 2020 to December 2023				
ASVI	0.111***	0.120***	0.936***	0.100**
	(0.040)	(0.040)	(0.160)	(0.040)
Log_Abnormal_Turnover	0.508***	0.514***	0.478***	0.415***
	(0.090)	(0.090)	(0.180)	(0.160)
Analyst_Coverage	-0.019***	0.012	-0.003	-0.017***
	(0.000)	(0.010)	(0.050)	(0.010)
Log_Size	0.127***	1.380***	8.640***	0.108***
	(0.040)	(0.120)	(0.580)	(0.040)
AdvertisingSales	-0.157	0.051	250.200	-0.189
	(0.640)	(2.060)	(197.360)	(0.550)
Log_IVOL	0.074	0.025	-0.082	0.087
	(0.060)	(0.080)	(0.140)	(0.090)
Idiosyncratic_Skewness	0.065*	0.052	-0.165***	0.011
	(0.030)	(0.030)	(0.050)	(0.040)
News_Coverage	0.042*	0.074**	0.035	0.038
	(0.020)	(0.030)	(0.060)	(0.020)
Earnings_Dummy	0.023	0.022	-0.361	0.076
	(0.130)	(0.120)	(0.580)	(0.130)
Lagged_ASVI				0.143**
				(0.070)
Lagged_Abnormal_Return				7.200
				(5.500)
_cons	1.920***	-20.700***	-26.300	2.050***
	(0.600)	(2.090)	(16.050)	(0.620)
Observations	26062	26062	26062	25920

Standard errors in parentheses, * $p < 10\%$, ** $p < 5\%$, *** $p < 1\%$

Table VI shows the results of several robustness checks in bps. Panel A contains the period from January 2019 to January 2020, Panel B contains the period from February 2020 to April 2020, and Panel C contains the period from May 2020 to December 2023.

The results from the full-period analysis show that ASVI is a significant predictor of Abnormal Return in the Nasdaq-100, and the regressions from each period disclose that the relationship varies, with the strongest effects observed during periods of higher market volatility. During the period ‘Before COVID-19’, ASVI exhibits significant predictive capability in the Pooled OLS, FE Panel Model, and GMM but lacks significance in the CCEMG model, which can be justified by the lower volatility and lack of attention triggers since the period was relatively stable compared to during and post-pandemic periods, making ASVI a less reliable tool for less volatile markets. Additionally, the CCEMG accounts for unobserved heterogeneity and common correlated effects across stocks, which are lower during stable periods, and therefore it is normal that the ASVI does not have significant results in this period.

In the following periods, ASVI has significant coefficients across all models. But the CCEMG model is the model that presents higher coefficients, underscoring the importance of addressing cross-sectional dependence in panel data. Since periods of uncertainty seem to amplify the influence of investor attention on Abnormal Returns. This finding highlights the importance of considering time-varying market factors since it suggests investors react more strongly to attention-grabbing events during volatile periods, which aligns with the literature that indicates markets tend to overreact to unexpected events (Bondt & Thaler, 1985).

The R-squared values suggest that ASVI and other investor attention proxies explain only a small portion of the variation in Abnormal Returns, consistent with the high noise in financial markets, but the low values can also be related to additional proxies that are not considered in this paper. Despite this, the model’s RMSE values indicate reasonable prediction accuracy, particularly during the pandemic, when volatility was at its peak. Additionally, the Arellano-Bond AR(2) test confirms the absence of second-order serial correlation, and the Hansen test ensures that the instruments used in the model are likely uncorrelated with the error term, supporting the robustness of the System GMM estimations. Finally, the Sargan test raises minor concerns post-COVID, that likely reflect structural changes in the market during this period, such as the widespread adoption of remote work, wars and fluctuating government policies, which may compromise the validity of the instruments used in the GMM model.

Hypothesis III: ASVI on Abnormal Return in Technology vs Non-Technology Stocks

This hypothesis goes beyond the paper from Da et al. (2011) by testing the impact of investor attention, measured by ASVI, on Abnormal Return (in bps), with a focus on technology versus non-technology stocks within the Nasdaq-100 because technology stocks tend to have higher volatility, more innovation cycles, and higher growth potential, which can lead to more investor attention and consequently higher returns (Cutter, 2024; Da et al., 2011; Moran, 2020; Padungsaksawasdi et al., 2019). Using the same four models as in the previous hypothesis, I find that ASVI has always positive coefficients, but it is not significant using the GMM model. This is justified by the fact that GMM assumes that the lagged instruments are not correlated with the error term, but this assumption may be violated due to the presence of strong cross-sectional dependence.

Looking at Table VII, there is a new variable, the ASVI Tech Dummy, which is used to account for the differential impact of ASVI on Abnormal Return in technology versus non-technology. The positive and significant coefficients from ASVI Tech Dummy in the Pooled OLS, FE Panel Model, and CCEMG models suggest that technology stocks exhibit a stronger relationship between ASVI and Abnormal Returns compared to non-technology stocks. However, since the magnitude is relatively small, the practical implications may not be substantial without further amplification from other factors. However, it is important to mention that this conclusion is specific to the Nasdaq-100, which is heavily weighted towards technology stocks, so the results may not generalise to indices with different compositions. Additionally, the period being studied is also peculiar so the results cannot be generalised to other periods without further studies. Although the R-squared values are lower than expected, the RMSE indicates that the models still achieve a level of prediction accuracy that aligns with the ASVI range. The Arellano-Bond tests confirm the validity of the dynamic panel model, with significant AR(1) at a 1% significance level, and non-significant AR(2), indicating no second-order serial correlation. The Hansen test validates the instrument set, confirming that the instruments are uncorrelated with the error term.

To check if the low ASVI values introduce noise in the data, it was done a robustness check in Table VIII by excluding 'noisy' firms that have low searches in Google because searches performed by a relatively low number of people are not accurate. After taking out these firms is possible to see that removing these 'noisy' firms hardly changes the results in bps but can introduce selection bias by excluding the firms with lower investor attention.

Table VII
Effect of ASVI on Abnormal Return in Tech vs Non-Tech Stocks

	(1) Pooled OLS	(2) FE Panel Model	(3) CCEMG	(4) System GMM
ASVI	0.097**	0.101**	0.521***	0.549
	(0.040)	(0.050)	(0.110)	(4.770)
ASVI_Tech_Dummy	0.129*	0.142*	0.180**	-0.021
	(0.070)	(0.070)	(0.070)	(1.390)
Log_Abnormal_Turnover	0.474***	0.502***	0.333**	-1.130
	(0.080)	(0.080)	(0.170)	(13.040)
Analyst_Coverage	-0.018***	0.006	0.023	-0.068
	(0.000)	(0.010)	(0.030)	(0.500)
Log_Size	0.133***	0.790***	6.710***	3.670
	(0.030)	(0.080)	(0.450)	(31.720)
AdvertisingSales	0.055	1.380	207.800	-23.700
	(0.510)	(1.580)	(169.860)	(225.780)
Log_IVOL	0.094*	0.030	-0.057	22.800
	(0.050)	(0.060)	(0.130)	(204.970)
Idiosyncratic_Skewness	0.038	0.037	-0.125***	-47.100
	(0.030)	(0.030)	(0.040)	(361.330)
News_Coverage	0.052***	0.054**	0.053	-0.754
	(0.020)	(0.30)	(0.050)	(8.700)
Earnings_Dummy	0.044	0.040	-0.320	0.731
	(0.110)	(0.110)	(0.480)	(5.980)
Lagged_ASVI				1.240
				(9.520)
Lagged_Abnormal_Return				13.200
				(80.250)
_cons	2.010***	-10.000***	-16.400	22.000
	(0.500)	(1.420)	(12.000)	(178.300)
Observations	33985	33985	33985	33842

Standard errors in parentheses, * $p < 10\%$, ** $p < 5\%$, *** $p < 1\%$

Table VIII

Effect of ASVI on Abnormal Return in Tech vs Non-Tech Stocks: Robustness

	(1) Pooled OLS	(2) FE Panel Model	(3) CCEMG	(4) System GMM
ASVI	0.165*	0.179**	0.542***	1.180
	(0.090)	(0.090)	(0.150)	(7.340)
ASVI_Tech_Dummy	0.266**	0.267**	0.256***	0.175
	(0.120)	(0.120)	(0.100)	(1.160)
Log_Abnormal_Turnover	0.534***	0.558***	0.343	0.160
	(0.090)	(0.090)	(0.210)	(11.660)
Analyst_Coverage	-0.021***	0.008	0.042	0.021
	(0.000)	(0.010)	(0.030)	(0.940)
Log_Size	0.116***	0.815***	6.080***	3.230
	(0.040)	(0.090)	(0.450)	(30.580)
AdvertisingSales	-0.194	3.740*	189.500	-15.000
	(0.570)	(1.990)	(158.090)	(57.740)
Log_IVOL	0.108*	0.092	0.045	18.800
	(0.060)	(0.080)	(0.110)	(183.380)
Idiosyncratic_Skewness	0.039	0.051	-0.130***	-25.800
	(0.030)	(0.030)	(0.050)	(94.260)
News_Coverage	0.068***	0.087***	0.052	-0.965
	(0.020)	(0.030)	(0.070)	(12.440)
Earnings_Dummy	-0.053	-0.055	-0.299	1.030
	(0.130)	(0.130)	(0.450)	(7.720)
Lagged_ASVI				2.260
				(18.530)
Lagged_Abnormal_Return				3.410
				(18.530)
_cons	2.400***	-10.400***	-18.200*	13.400
	(0.560)	(1.590)	(10.250)	(92.510)
Observations	24806	24806	24806	24702

Standard errors in parentheses, * $p < 10\%$, ** $p < 5\%$, *** $p < 1\%$

V. Discussion

This thesis findings reveal that ASVI has predictive capability over Abnormal Returns in the short term, consistent with studies from Da et al. (2011), Barber and Odean (2008), Tan and Tas (2019), and Ayaz et al. (2021). Even though it focuses on capturing retail investor attention, and retail volume only accounts for 20% of stock market activity (Eaton et al., 2022). This paper particularly studies the relationship between ASVI and Abnormal Returns in highly volatile market conditions, such as during and post-COVID-19 pandemic, where the predictive capability is even stronger. These results provide valuable insights into investor behaviour during periods of high uncertainty since existing literature focuses only on general market conditions, and this thesis uses firm-specific keywords to further study investor attention. However, since this thesis studies a particular period, the applicability to more stable conditions require careful consideration. Looking at the numbers, I conclude that before COVID-19, ASVI is not significant under CCEMG, which suggests that ASVI is a less reliable tool for less volatile markets. This result changes in the following two periods due to strong cross-sectional dependence caused by the COVID-19 pandemic, since in these periods all the models are significant. The variable Log Abnormal Turnover also became significant during and post-COVID-19, which shows the importance of trading activity in predicting Abnormal Returns during these periods. So, I can conclude that variables like ASVI and Log Abnormal Turnover are more robust across models because they capture investor behaviour more directly. At the same time, the variable News Coverage measures the numbers of news related to each firm in Refinitiv so the lack of significance in the CCEMG model may be due to the aggregation effects, as firm-specific news impacts are not uniformly correlated across groups. Second, I study the sector-specific impact of ASVI and conclude that its influence on Abnormal Return is stronger for technology stocks compared to non-technology stocks, which can be justified by the higher volatility and growth potential (Cutter, 2024; Moran, 2020). Nevertheless, these conclusions have to be carefully generalised since this study exclusively focuses on the Nasdaq-100, and the results may not generalise to indices with different compositions, like less tech-heavy indices. Additionally, the particular market conditions of this period can also have a stronger impact on the results, effects that may not happen under different conditions. In this hypothesis, the ASVI is significant across all models except the GMM, which can indicate the difficulty of applying this method in datasets with strong cross-sectional dependence.

Regarding the methodologies, I primarily rely on the CCEMG model due to its significance related to the strong cross-sectional dependence and heterogeneity of the data. The Pooled OLS assumes homogeneity across firms and periods, so it cannot account for unobserved heterogeneity or cross-sectional dependence, which biases the data because the firms have unique characteristics and are exposed to the same macroeconomic factors. This means that even though some variables present significant coefficients these can happen only because the model captures their raw relationships with Abnormal Return, which overstates the effects. The FE Panel Model already accounts for the time-invariant firm-specific characteristics and time-specific shocks, so it already eliminates the bias from unobserved differences between firms. Lastly, the GMM addresses endogeneity by including lagged variables, but its assumptions that residuals are uncorrelated across firms make it less reliable under strong cross-sectional dependence, as seen with the lack of significant results. Therefore, this model is not the most adequate for datasets with high volatility. Finally, I conclude that ASVI effectiveness may vary depending on the econometric method, and no single method fully addresses all econometric challenges, since by focusing on CCEMG, I am not fully considering endogeneity, autocorrelation, time-varying factors and heteroskedasticity, which bias the results.

Furthermore, the study has several limitations regarding data retrieval. First, despite efforts to include delisted companies, Refinitiv's lack of data for merged and acquired companies introduces potential survivorship bias, which distorts the results if these firms had unique investor attention patterns. Second, News Coverage is limited to Refinitiv's news based on the tickers, which results in the underrepresentation of attention-driven effects of certain firms. Third, Analyst Coverage and Advertising lack weekly data, which reduces the precision of these control variables and might weaken the results' robustness. Fourth, the SVI data is also constrained by Google Trends' relative measures and its five-year restriction, which excludes searches by a relatively low number of users and makes it difficult to study the long-term effect. Finally, while institutional investors are familiar with Google, they rarely use it for financial research. As a result, Google Trends only reflects searches that correspond to around 20% of the stock market activity (Eaton et al., 2022), potentially biasing the results since they do not account for institutional investor attention, which represents the majority of market activity.

This study can be further developed in future research by extending the analysis to other attention-grabbing sectors, and by analysing the impact of firm size, since the variable Log Size is positive and significant across several models and periods, which indicates that the firm size tends to attract more investor interest and consequently more Abnormal Returns. Further research can also examine the possibility of reversal effects over longer timeframes, which was not possible due to the lack of data. Finally, it can also analyse the profitability of the findings to assess their practical application in the stock market.

VI. Conclusion

The Google Search Volume Index is an accessible dataset that can be exploited even by investors without advanced financial literacy. Therefore, it is a good proxy for investor attention that can be used to predict Abnormal Returns.

I find that ASVI is correlated but distinct from other investor attention proxies, so ASVI can be used to further explore its impact. I also find that an increase in ASVI predicts positive Abnormal Return in the short term, as in Da et al. (2011), Barber & Odean (2008), Tan & Taş (2019), and Ayaz et al. (2021), but instead of studying the price reversal theory, I focus on the impact of market-wide factors such as the COVID-19 that enables the conclusion that ASVI has more positive impact on Abnormal Return when the market is more volatile and uncertainty, which aligns with the literature that states that market overreacts to unexpected events (Bondt & Thaler, 1985). Additionally, I also observe that ASVI's effect is more positively pronounced in technology stocks compared to non-technology stocks possibly due to the higher investor attention to this sector, which can be explained by the higher opportunities to achieve Abnormal Returns due to the fact that these stocks are recognised by their tech innovations, higher growth potential and higher market volatility, that grab investor attention (Cutter, 2024; Moran, 2020).

Therefore, these results underscore ASVI's purpose as a valuable tool for investors, particularly for retail investors, who seek to make more informed decisions in a financial environment characterised by extensive amounts of information to analyse, by providing evidence of a direct measure of investor attention. However, the magnitude of the impact that ASVI has in the short-term Abnormal Return is very small, so I can conclude that this tool should be used as one of many tools rather than a distinctive predictor.

These findings support behavioural finance theories that highlight the importance of attention biases, which means they also challenge the Efficient Market Hypothesis, which states that all available information is already priced in by the market (Fama, 1970; Peng & Xiong, 2006). With this in mind, ASVI could possibly be used to draw trading strategies that take advantage of market inefficiencies, however, to do this further research would be needed as among others, it is important to consider factors such as trading costs that could overwhelm any possible profits.

VII. References

Akinwande, M. O., Dikko, H. G., & Samson, A. (2015). Variance Inflation Factor: As a Condition for the Inclusion of Suppressor Variable(s) in Regression Analysis. *Open Journal of Statistics*, *5*(7), 754-767. http://dx.doi.org/10.4236/ojs.2015.57075

Angrisani, M., Burke, J., Lusardi, A., & Mottola, G. (2023). The evolution of financial literacy over time and its predictive power for financial outcomes: evidence from longitudinal data. *Journal of Pension Economics and Finance*, *22*(4), 640–657. https://doi.org/10.1017/S1474747222000154

Aouadi, A., Arouri, M., & Teulon, F. (2013). Investor attention and stock market activity: Evidence from France. *Economic Modelling*, *35*, 674–681. https://doi.org/10.1016/j.econmod.2013.08.034

Arellano, M. (1991). Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations. *The Review of Economic Studies*, *58*(2), 277-297. https://doi.org/10.2307/2297968

Arellano, M., & Bover, O. (1995). Another look at the instrumental variable estimation of error-components models. *Journal of Econometrics*, *68*(1), 29–51. https://doi.org/10.1016/0304-4076(94)01642-D

Askitas, N., & Zimmermann, K. F. (2009). Google Econometrics and Unemployment Forecasting. *Applied Economics Quarterly*, *55*(2), 107–120. https://doi.org/10.2139/ssrn.1480251

Ayala, M. J., Gonzálvez-Gallego, N., & Arteaga-Sánchez, R. (2024). Google search volume index and investor attention in stock market: a systematic review. *Financial Innovation*, *10*(1), 1-29. https://doi.org/10.1186/s40854-023-00606-y

Ayaz, B., Ullah, H., Khan, M. K., & Jan, S. (2021). The Effect of Google Search Volume Index on the Stock Market Excess Returns. Evidence from Listed firms in Pakistan stock Exchange. *Review of Education, Administration & LAW*, *4*(1), 23–35. https://doi.org/10.47067/real.v4i1.108

Baltagi, B. H. (2008). *Econometric Analysis of Panel Data* (3rd ed.). John Wiley & Sons, Ltd.

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), 239–264. 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

Basuony, M. A. K., Bouaddi, M., Ali, H., & EmadEldeen, R. (2021). The effect of COVID-19 pandemic on global stock markets: Return, volatility, and bad state probability dynamics. *Journal of Public Affairs*, *22*(4). https://doi.org/10.1002/pa.2761

Ben-Rephael, A., Da, Z., & Israelsen, R. (2017). It depends on where you search: Institutional investor attention and underreaction to news. *Review of Financial Studies, 30*(9), 3009-3047. https://doi.org/10.1093/rfs/hhx031

Bernard, V. L. (1987). Cross-Sectional Dependence and Problems in Inference in Market-Based Accounting Research. *Journal of Accounting Research*, *25*(1), 1–48. https://doi.org/10.2307/2491257

Bianchi, T. (2024, February). *Market share of leading search engines worldwide from January 2015 to January 2024*. Statista. https://www.statista.com/statistics/1381664/worldwide-all-devices-market-share-of-search-engines/

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

Bondt, W. F. M. D., & Thaler, R. (1985). Does the Stock Market Overreact? *The Journal of Finance*, *40*(3), 793–805. https://doi.org/10.1111/j.1540-6261.1985.tb05004.x

Cevik, E., Altinkeski, B. K., Cevik, E. I., & Dibooglu, S. (2022). Investor sentiments and stock markets during the COVID-19 pandemic. *Financial Innovation*, *8*(1), 1-34. https://doi.org/10.1186/s40854-022-00375-0

Chundakkadan, R., & Nedumparambil, E. (2022). In search of COVID-19 and stock market behavior. *Global Finance Journal*, *54, 100639*. https://doi.org/10.1016/j.gfj.2021.100639

Cutter, P. (2024, October). *Navigating tech volatility*. Schroders. https://www.schroders.com/en-us/us/intermediary/insights/navigating-tech-volatility/

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

Da, Z., Engelberg, J., & Gao, P. (2015). The sum of all FEARS investor sentiment and asset prices. *Review of Financial Studies*, *28*(1), 1–32. https://doi.org/10.1093/rfs/hhu072

Das, S. R., & Chen, M. Y. (2007). Yahoo! for Amazon: Sentiment Extraction from Small Talk on the Web. *Management Science*, *53*(9), 1375–1388. https://doi.org/10.1287/mnsc.1070.0704

Ding, R., & Hou, W. (2015). Retail investor attention and stock liquidity. *Journal of International Financial Markets, Institutions and Money*, *37*, 12–26. https://doi.org/10.1016/j.intfin.2015.04.001

Ditzen, J. (2018). Estimating Dynamic Common-Correlated Effects in Stata. *The Stata Journal*, *18*(3), 585–617. https://doi.org/10.1177/1536867X1801800306

Drake, M. S., Roulstone, D. T., & Thornock, J. R. (2012). Investor Information Demand: Evidence from Google Searches Around Earnings Announcements. *Journal of Accounting Research*, *50*(4), 1001–1040. https://doi.org/10.1111/j.1475-679X.2012.00443.x

Eaton, G. W., Green, T. C., Roseman, B. S., & Wu, Y. (2022). Retail trader sophistication and stock market quality: Evidence from brokerage outages. *Journal of Financial Economics*, *146*(2), 502–528. https://doi.org/10.1016/j.jfineco.2022.08.002

Fama, E. F. (1970). Efficient Capital Markets: A Review of Theory and Empirical Work. *The Journal of Finance*, *25*(2), 383-417. https://doi.org/10.2307/2325486

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, *33*(1), 3–56. https://doi.org/10.1016/0304-405X(93)90023-5

Fama, E. F., & MacBeth, J. D. (1973). Risk, Return, and Equilibrium: Empirical Tests. *Journal of Political Economy*, *81*(3), 607–636. https://doi.org/10.1086/260061

Fang, L., & Peress, J. (2009). Media Coverage and the Cross-section of Stock Returns. *The Journal of Finance*, *64*(5), 2023–2052. https://doi.org/10.1111/j.1540-6261.2009.01493.x

Gervais, S., Kaniel, R., & Mingelgrin, D. H. (2001). The High-Volume Return Premium. *The Journal of Finance*, *56*(3), 877–919. https://doi.org/10.1111/0022-1082.00349

Ginsberg, J., Mohebbi, M. H., Patel, R. S., Brammer, L., Smolinski, M. S., & Brilliant, L. (2009). Detecting influenza epidemics using search engine query data. *Nature*, *457*, 1012–1014. https://doi.org/10.1038/nature07634

Google. (2024). *Google Trends*. Google. https://trends.google.com/trends/

Gow, I. D., Ormazabal, G., & Taylor, D. J. (2010). Correcting for Cross-Sectional and Time-Series Dependence in Accounting Research. *The Accounting Review*, *85*(2), 483–512. https://doi.org/10.2308/accr.2010.85.2.483

Grullon, G., Kanatas, G., & Weston, J. P. (2004). Advertising, Breadth of Ownership, and Liquidity. *Review of Financial Studies*, *17*(2), 439–461. https://doi.org/10.1093/rfs/hhg039

Hausman, J. A. (1978). Specification Tests in Econometrics. *Econometrica*, *46*(6), 1251–1271. https://doi.org/10.2307/1913827

Hirshleifer, D., & Teoh, S. H. (2003). Limited attention, information disclosure, and financial reporting. *Journal of Accounting and Economics*, *36*(1–3), 337–386. https://doi.org/10.1016/j.jacceco.2003.10.002

Kayani, U. N., Aysan, A. F., Khan, M., Khan, M., Mumtaz, R., & Irfan, M. (2024). Unleashing the pandemic volatility: A glimpse into the stock market performance of developed economies during COVID-19. *Heliyon*, *10*(4). https://doi.org/10.1016/j.heliyon.2024.e25202

Kumar, A. (2009). Who Gambles in the Stock Market? *The Journal of Finance*, *64*(4), 1889–1933. https://doi.org/10.1111/j.1540-6261.2009.01483.x

Lobão, J., Pacheco, L., & Pereira, C. (2017). The use of the recognition heuristic as an investment strategy in European stock markets. *Journal of Economics, Finance and Administrative Science*, *22*(43), 207–223. https://doi.org/10.1108/JEFAS-01-2017-0013

Lou, D. (2014). Attracting Investor Attention through Advertising. *The Review of Financial Studies*, *27*(6), 1797–1829. https://doi.org/10.1093/rfs/hhu019

Merton, R. C. (1987). A Simple Model of Capital Market Equilibrium with Incomplete Information. *The Journal of Finance*, *42*(3), 483–510. https://doi.org/10.1111/j.1540-6261.1987.tb04565.x

Mondria, J., Wu, T., & Zhang, Y. (2010). The determinants of international investment and attention allocation: Using internet search query data. *Journal of International Economics*, *82*(1), 85–95. https://doi.org/10.1016/j.jinteco.2010.04.007

Moran, M. (2020, January). *Performance and Volatility for Sectors in the 2010s*. S&P Global. https://www.spglobal.com/en/research-insights/market-insights/performance-and-volatility-for-sectors-in-the-2010s

Newey, W. K., & West, K. D. (1987). A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. *Econometrica*, *55*(3), 703-708. https://doi.org/10.2307/1913610

Padungsaksawasdi, C., Treepongkaruna, S., & Brooks, R. (2019). Investor attention and stock market activities: New evidence from panel data. *International Journal of Financial Studies*, *7*(2), 1-19. https://doi.org/10.3390/ijfs7020030

Peng, L. (2005). Learning with Information Capacity Constraints. *Journal of Financial and Quantitative Analysis*, *40*(2), 307–329. https://doi.org/10.1017/S0022109000002325

Peng, L., & Xiong, W. (2006). Investor attention, overconfidence and category learning. *Journal of Financial Economics*, *80*(3), 563–602. https://doi.org/10.1016/j.jfineco.2005.05.003

Pesaran, M. H. (2006). Estimation and Inference in Large Heterogeneous Panels with a Multifactor Error Structure. *Econometrica*, *74*(4), 967–1012. https://doi.org/10.1111/j.1468-0262.2006.00692.x

Preis, T., Reith, D., & Stanley, H. E. (2010). Complex dynamics of our economic life on different scales: Insights from search engine query data. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *368*(1933), 5707–5719. https://doi.org/10.1098/rsta.2010.0284

Siganos, A., Vagenas-Nanos, E., & Verwijmeren, P. (2014). Facebook's daily sentiment and international stock markets. *Journal of Economic Behavior and Organization*, *107*, 730–743. https://doi.org/10.1016/j.jebo.2014.06.004

Simon, H. A. (1971). Designing Organizations For An Information-Rich World. In M. Greenberger (Ed.), *Computers, Communications, and the Public Interest* (pp. 37–72). The Johns Hopkins Press.

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

Statista Research Department. (2023, November). *Global online trading market from 2020 with forecasts to 2026*. Statista. https://www.statista.com/statistics/1260026/forecast-global-online-trading-platform-market/

Tan, S. D., & Taş, O. (2019). Investor attention and stock returns: Evidence from Borsa Istanbul. *Borsa Istanbul Review*, *19*(2), 106–116. https://doi.org/10.1016/j.bir.2018.10.003

Tetlock, P. C. (2007). Giving Content to Investor Sentiment: The Role of Media in the Stock Market. *The Journal of Finance*, *62*(3), 1139–1168. https://doi.org/10.1111/j.1540-6261.2007.01232.x

Thaler, R. H. (2005). Advances in Behavioral Finance. In *Princeton University press* (Vol. 2). Journal of Finance. https://doi.org/10.2307/2329257

Thompson II, R. B., Olsen, C., & Dietrich, J. R. (1987). Attributes of News about Firms: An Analysis of Firm-Specific News Reported in the Wall Street Journal Index. *Journal of Accounting Research*, *25*(2), 245–274. https://doi.org/10.2307/2491017

Ullah, S., Khan, S., Hashmi, N. I., & Alam, M. S. (2023). COVID-19 pandemic and financial market volatility: A quantile regression approach. *Heliyon*, *9*(10). https://doi.org/10.1016/j.heliyon.2023.e21131

Varian, H. R., & Choi, H. (2012). Predicting the Present with Google Trends. *Economic Record*, *88*(s1), 2–9. https://doi.org/10.1111/j.1475-4932.2012.00809.x

Wheat, C., & Eckerd, G. (2024, November). *The changing demographics of retail investors*. JPMorganChase. https://www.jpmorganchase.com/institute/all-topics/financial-health-wealth-creation/the-changing-demographics-of-retail-investors

Yahya, F., Shaohua, Z., Abbas, U., & Waqas, M. (2021). COVID-19-induced Returns, Attention, Sentiments and Social Isolation: Evidence from Dynamic Panel Model. *Global Business Review*, 1–19. https://doi.org/10.1177/0972150921996174

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

VIII. Appendices

Appendix I. Companies Description

Company Name	Search Term Used	Ticker Symbol	Sector
Amgen Inc	Amgen	AMGN	Non-Tech
Apple Inc	Apple	AAPL	Tech
Applied Materials Inc	Applied Materials	AMAT	Tech
Intel Corp	Intel	INTC	Tech
Paccar Inc	Paccar	PCAR	Non-Tech
Microsoft Corp	Microsoft	MSFT	Tech
Xilinx Inc	Xilinx	XLNX	Tech
Intuit Inc	Intuit	INTU	Tech
Willis Towers Watson Plc	Willis Towers Watson	WTW	Non-Tech
Starbucks Corp	Starbucks	SBUX	Non-Tech
Fiserv Inc	Fiserv	FI	Tech
Citrix Systems Inc	Citrix Systems	CTXS	Tech
Amazon.Com Inc	Amazon	AMZN	Tech
Ebay Inc	Ebay	EBAY	Non-Tech
Biogen Inc	Biogen	BIIB	Non-Tech
Check Point Software Techn	Check Point	CHKP	Tech
Gilead Sciences Inc	Gilead Sciences	GILD	Non-Tech
Symantec/ Gen Digital Inc	Symantec/Nortonlifelock/Gen Digital	GEN	Tech
Comcast Corp	Comcast	CMCSA	Tech
Fastenal Co	Fastenal	FAST	Non-Tech
Autodesk Inc	Autodesk	ADSK	Tech
Cognizant Tech Solutions	Cognizant	CTSH	Tech
Celgene Corp	Celgene	CELG	Non-Tech
Nvidia Corp	Nvidia	NVDA	Tech
Alphabet Inc	Alphabet	GOOGL	Tech
Intuitive Surgical Inc	Intuitive Surgical	ISRG	Non-Tech
Vertex Pharmaceuticals Inc	Vertex Pharmaceuticals	VRTX	Non-Tech
Henry Schein Inc	Henry Schein	HSIC	Non-Tech
Baidu Inc	Baidu	BIDU	Tech
Activision Blizzard Inc	Activision Blizzard	ATVI	Tech
Automatic Data Processing	Automatic Data Processing	ADP	Tech
Ross Stores Inc	Ross Stores	ROST	Non-Tech
O'reilly Automotive Inc	O'reilly	ORLY	Non-Tech
Twenty-First Century Fox Inc/Fox Corp	Fox	FOXA	Non-Tech
Cerner Corp	Cerner	CERN	Tech
Booking Holdings Inc	Booking Holdings	BKNG	Non-Tech
Viatris Inc	Viatris	VTRS	Non-Tech
Micron Technology Inc	Micron Technology	MU	Tech
Dollar Tree Inc	Dollar Tree	DLTR	Non-Tech
Alexion Pharmaceuticals Inc	Alexion Pharmaceuticals	ALXN	Non-Tech
Sirius Xm Holdings Inc	Sirius Xm	SIRI	Non-Tech
Monster Beverage Corp	Monster Beverage	MNST	Non-Tech
Broadcom Inc	Broadcom	AVGO	Tech
Texas Instruments Inc	Texas Instruments	TXN	Tech
Mondelez International Inc	Mondelez	MDLZ	Non-Tech
Paychex Inc	Paychex	PAYX	Non-Tech
Facebook/ Meta Platforms Inc	Facebook/Meta	META	Tech
Analog Devices Inc	Analog Devices	ADI	Tech
Regeneron Pharmaceuticals	Regeneron Pharmaceuticals	REGN	Non-Tech
Verisk Analytics Inc	Verisk Analytics	VRSK	Non-Tech
Western Digital Corp	Western Digital	WDC	Tech
Liberty Global Ltd	Liberty Global	LBTYK	Non-Tech
Netflix Inc	Netflix	NFLX	Tech
Tesla Inc	Tesla	TSLA	Tech
Charter Communications Inc	Charter Communications	CHTR	Non-Tech
Adobe Inc	Adobe	ADBE	Tech
Marriott Intl Inc	Marriott	MAR	Non-Tech
Illumina Inc	Illumina	ILMN	Non-Tech
Lam Research Corp	Lam Research	LRCX	Tech
Electronic Arts Inc	Electronic Arts	EA	Tech
American Airlines Group Inc	American Airlines	AAL	Non-Tech
Walgreens Boots Alliance Inc	Walgreens Boots Alliance	WBA	Non-Tech
Kraft Heinz Co	Kraft Heinz	KHC	Non-Tech
Biomarin Pharmaceutical Inc	Biomarin Pharmaceutical	BMRN	Non-Tech
Jd.Com Inc	Jd	JD	Tech
Cisco Systems Inc	Cisco Systems	CSCO	Tech

Skyworks Solutions Inc	Skyworks Solutions	SWKS	Tech
Incyte Corp	Incyte	INCY	Non-Tech
Paypal Holdings Inc	Paypal	PYPL	Tech
T-Mobile Us Inc	T-Mobile	TMUS	Non-Tech
Expedia Group Inc	Expedia	EXPE	Non-Tech
Ulta Beauty Inc	Ulta Beauty	ULTA	Non-Tech
Maxim Integrated Products	Maxim	MXIM	Tech
Trip Com Group Ltd	Ctrip Stock/Trip Com Group	TCOM	Non-Tech
Csx Corp	Csx	CSX	Non-Tech
Qualcomm Inc	Qualcomm	QCOM	Tech
Netease Inc	Netease	NTES	Tech
Microchip Technology Inc	Microchip	MCHP	Tech
Cintas Corp	Cintas	CTAS	Non-Tech
Kla Corp	Kla	KLAC	Tech
Hasbro Inc	Hasbro	HAS	Non-Tech
Hunt (Jb) Transprt Svcs Inc	Jb Hunt	JBHT	Non-Tech
Idexx Labs Inc	Idexx Labs	IDXX	Non-Tech
Wynn Resorts Ltd	Wynn Resorts	WYNN	Non-Tech
Mercadolibre Inc	Mercadolibre	MELI	Tech
Align Technology Inc	Align Technology	ALGN	Non-Tech
Costco Wholesale Corp	Costco Wholesale	COST	Non-Tech
Cadence Design Systems Inc	Cadence	CDNS	Tech
Workday Inc	Workday	WDAY	Tech
Synopsys Inc	Synopsys	SNPS	Tech
Asml Holding Nv	Asml Holding	ASML	Tech
Take-Two Interactive Sftwr	Take-Two Interactive	TTWO	Tech
Pepsico Inc	Pepsico	PEP	Non-Tech
Nxp Semiconductors Nv	Nxp Semiconductors	NXPI	Tech
Xcel Energy Inc	Xcel Energy	XEL	Non-Tech
Advanced Micro Devices	Advanced Micro Devices	AMD	Tech
Lululemon Athletica Inc	Lululemon Athletica	LULU	Non-Tech
United Airlines Holdings Inc	United Airlines	UAL	Non-Tech
Netapp Inc	Netapp	NTAP	Tech
Verisign Inc	Verisign	VRSN	Tech
Exelon Corp	Exelon	EXC	Non-Tech
Copart Inc	Copart	CPRT	Non-Tech
Ansys Inc	Ansys	ANSS	Tech
Costar Group Inc	Costar	CSGP	Tech
Cdw Corp	Cdw	CDW	Tech
Splunk Inc	Splunk	SPLK	Tech
Seattle Genetics/Seagen Inc	Seattle Genetics/Seagen	SGEN	Non-Tech
Dexcom Inc	Dexcom	DXCM	Non-Tech
Zoom Video Comunications Inc	Zoom	ZM	Tech
Docusign Inc	Docusign	DOCU	Tech
Moderna Inc	Moderna	MRNA	Non-Tech
Pdd Holdings Inc	Pdd	PDD	Tech
Keurig Dr Pepper Inc	Keurig Dr Pepper	KDP	Non-Tech
American Electric Power Co	American Electric Power	AEP	Non-Tech
Atlassian Corp	Atlassian	TEAM	Tech
Marvell Technology Inc	Marvell Technology	MRVL	Tech
Peloton Interactive Inc	Peloton	PTON	Non-Tech
Okta Inc	Okta	OKTA	Tech
Match Group Inc	Match Group	MTCH	Tech
Honeywell International Inc	Honeywell	HON	Non-Tech
Crowdstrike Holdings Inc	Crowdstrike	CRWD	Tech
Zscaler Inc	Zscaler	ZS	Tech
Datadog Inc	Datadog	DDOG	Tech
Airbnb Inc	Airbnb	ABNB	Non-Tech
Palo Alto Networks Inc	Palo Alto Networks	PANW	Tech
Fortinet Inc	Fortinet	FTNT	Tech
Lucid Group Inc	Lucid Group	LCID	Tech
Old Dominion Freight	Old Dominion	ODFL	Non-Tech
Constellation Ene Corp	Constellation	CEG	Non-Tech
Astrazeneca Plc	Astrazeneca	AZN	Non-Tech
Enphase Energy Inc	Enphase Energy	ENPH	Tech
Baker Hughes Co	Baker Hughes	BKR	Non-Tech
Globalfoundries Inc	Globalfoundries	GFS	Tech
Warner Bros Discovery Inc	Warner Bros	WBD	Non-Tech
Diamondback Energy Inc	Diamondback Energy	FANG	Non-Tech
Rivian Automotive Inc	Rivian Automotive	RIVN	Tech
Ge Healthcare Technologi Inc	Ge Healthcare	GEHC	Non-Tech
On Semiconductor Corp	On Semiconductor	ON	Tech
Trade Desk Inc	Trade Desk	TTD	Tech
Coca-Cola Europacific Partne	Coca-Cola Europacific Partne	KO	Non-Tech
Roper Technologies Inc	Roper	ROP	Tech
Mongodb Inc	Mongodb	MDB	Tech
Doordash Inc	Doordash	DASH	Non-Tech

Appendix II. Correlation Matrix (01/2019 – 01/2020)

	Abnormal Return	ASVI	Log Abnormal Turnover	Log Size	Analyst Coverage	Advertising/Sales	Log IVOL	Idiosyncratic Skewness	News Coverage
Abnormal Return	1.000								
ASVI	0.036	1.000							
Log Abnormal Turnover	-0.009	0.078	1.000						
Log Size	0.036	-0.010	-0.027	1.000					
Analyst Coverage	-0.001	-0.006	-0.022	0.457	1.000				
Advertising/Sales	0.003	-0.011	-0.003	-0.149	-0.010	1.000			
Log IVOL	-0.014	-0.028	-0.006	-0.457	-0.118	0.167	1.000		
Idiosyncratic Skewness	-0.002	-0.005	-0.022	-0.065	-0.111	0.018	-0.010	1.000	
News Coverage	0.031	0.030	0.201	0.432	0.385	-0.041	-0.089	-0.042	1.000

Appendix III. Correlation Matrix (02/2020 – 04/2020)

	Abnormal Return	ASVI	Log Abnormal Turnover	Log Size	Analyst Coverage	Advertising/Sales	Log IVOL	Idiosyncratic Skewness	News Coverage
Abnormal Return	1.000								
ASVI	0.057	1.000							
Log Abnormal Turnover	0.107	0.229	1.000						
Log Size	0.018	-0.015	-0.009	1.000					
Analyst Coverage	-0.014	-0.010	0.025	0.494	1.000				
Advertising/Sales	0.013	-0.035	-0.017	-0.152	-0.109	1.000			
Log IVOL	0.051	-0.113	-0.109	-0.373	-0.179	0.138	1.000		
Idiosyncratic Skewness	-0.018	-0.017	-0.008	0.074	0.056	-0.050	0.080	1.000	
News Coverage	0.070	-0.050	0.223	0.406	-0.337	-0.009	0.003	0.075	1.000

Appendix IV. Correlation Matrix (05/2020 – 12/2023)

	Abnormal Return	ASVI	Log Abnormal Turnover	Log Size	Analyst Coverage	Advertising/Sales	Log IVOL	Idiosyncratic Skewness	News Coverage
Abnormal Return	1.000								
ASVI	0.020	1.000							
Log Abnormal Turnover	0.039	0.092	1.000						
Log Size	0.015	0.006	-0.016	1.000					
Analyst Coverage	-0.024	-0.006	-0.008	0.397	1.000				
Advertising/Sales	-0.006	0.007	0.012	-0.118	-0.064	1.000			
Log IVOL	0.007	-0.056	-0.047	-0.218	-0.155	0.181	1.000		
Idiosyncratic Skewness	0.011	-0.007	-0.033	0.007	-0.015	0.013	0.022	1.000	
News Coverage	0.027	0.058	0.198	0.430	0.233	0.015	0.039	0.014	1.000