

DOCTORAL THESIS



UCAM

UNIVERSIDAD CATÓLICA
DE MURCIA

INTERNATIONAL DOCTORAL SCHOOL

Doctoral Programme in Social Sciences

Investor Attention and Stock Market Behaviour in Europe:
Evidence from Google search data

Author:

María José Ayala Marín

Supervisors:

Dr Rocío Arteaga Sánchez

Dr Nicolás González Gallego

Murcia, July 2025

DOCTORAL THESIS



UCAM

UNIVERSIDAD CATÓLICA
DE MURCIA

INTERNATIONAL DOCTORAL SCHOOL

Doctoral Programme in Social Sciences

Investor Attention and Stock Market Behaviour in Europe:
Evidence from Google search data

Author:

María José Ayala Marín

Supervisors:

Dr Rocío Arteaga Sánchez

Dr Nicolás González Gallego

Murcia, July 2025

THESIS SUPERVISORS' AUTHORISATION FOR THESIS SUBMISSION

Prof. Dr. Rocío Arteaga Sánchez and Prof. Dr. Nicolás González Gallego, as Supervisors⁽¹⁾ of the Doctoral Thesis "Investor Attention and Stock Market Behaviour in Europe: Evidence from Google search data" by Ms María José Ayala Marín in the Doctorate Programme of Social Sciences, **authorise its submission**, given that it meets the required conditions for its defence.

Which I hereby sign in compliance with Spanish Royal Decree 99/2011, of 28 January, in Murcia, on 14th July 2025

**ARTEAGA
SANCHEZ
ROCIO -
48913578F**

Firmado
digitalmente por
ARTEAGA
SANCHEZ ROCIO -
48913578F
Fecha: 2025.07.14
13:55:53 +02'00'

**NICOLAS
GONZALVEZ
Z GALLEGU
- 48612261J**

Firmado
digitalmente por
NICOLAS
GONZALVEZ
GALLEGO -
48612261J
Fecha: 2025.07.14
14:06:02 +02'00'

⁽¹⁾ If the Thesis is supervised by more than one Supervisor, both must be mentioned and both must sign.

ABSTRACT

This thesis contributes to the growing literature on investor attention by reinforcing the usefulness of internet search data as a powerful tool for both forecasting and nowcasting stock market performance. Specifically, it presents an empirical analysis that applies both nowcasting and forecasting methodologies using aggregated Google search data through the Google Search Volume Index (GSVI) across European stock markets. To provide a solid foundation for the research, the thesis includes a comprehensive literature review and develops a theoretical framework that incorporates the primary theories on investor behaviour and informational demand.

The GSVI reflects the real-time demand for information by investors and, consequently, their intentions and behavioural responses to market stimuli. The central objective of this thesis is to assess the capacity of the GSVI to anticipate and predict movements in the European stock market. To achieve this, three pairs of hypotheses are proposed to analyze, through panel data regression models, the capacity to estimate and forecast three key financial metrics: stock return, volatility, and trading volume.

The empirical analysis is conducted using panel data from 52 publicly traded companies listed in the European stock market and included in the EUROSTOXX 50 index over the period 2014–2024. This cross-country sample enhances the relevance of the findings by offering a broader, more generalizable perspective beyond country-specific stock indices. Three distinct regression models are estimated to determine the influence of the GSVI on returns, volatility, and trading volume.

The results from the forecasting models show that aggregated online search behaviour holds significant predictive power, particularly over trading volume and volatility. Moreover, this predictive capacity increases as the forecasting horizon extends. Thus, the thesis concludes that early signals of investor attention reinforce their value in financial forecasting and, consequently, in investment analysis and decision-making.

Building on these results, the thesis validates the GSVI as a reliable proxy for investor attention in European stock markets, supporting the integration non-traditional data in behavioural finance models. In particular, concerning nowcasting, while its capacity to reflect an immediate perception of financial market conditions is limited, it proves valuable as a complementary indicator for anticipating very short-term changes in returns, volatility, and trading volume.

Moreover, the thesis underscores the importance of decisions related to the construction of the aggregated search index. It offers relevant insights, showing that alternative methods of building the GSVI are aligned with the financial variables analysed and the different forecasting horizons. These conclusions have practical implications for portfolio risk management, sentiment-driven trading strategies, and the development of more effective market signals. Finally, the thesis outlines future research directions to further enhance the use of digital attention metrics in financial markets.

Keywords: Investor attention, Behavioural finance, investment theory, data analysis, Google Search Volume Index (GSVI), European stock markets.

RESUMEN

Esta tesis contribuye a la creciente literatura sobre la atención de los inversores, reforzando la utilidad de los datos de búsquedas en internet como una herramienta poderosa tanto para la predicción (*forecasting*) como para la estimación en tiempo real (*nowcasting*) del comportamiento del mercado bursátil. En concreto, presenta un análisis empírico que aplica metodologías de *nowcasting* y *forecasting* utilizando datos agregados de búsqueda en Google a través de Google Search Volume Index (GSVI) en los mercados bursátiles europeos. Para proporcionar una base sólida a la investigación, la tesis incluye una revisión exhaustiva de la literatura y desarrolla un marco teórico que incorpora las principales teorías sobre el comportamiento de los inversores y la demanda de información.

El índice GSVI refleja la demanda de información en tiempo real por parte de los inversores y, en consecuencia, sus intenciones y respuestas conductuales ante estímulos del mercado. El objetivo principal de esta tesis es evaluar la capacidad del GSVI para anticipar y predecir movimientos en el mercado bursátil europeo. Para ello, se proponen tres pares de hipótesis para analizar, a través de modelos de regresión con datos de panel, la capacidad para estimar y predecir tres métricas financieras claves: rentabilidad, volatilidad y volumen de negociación.

El análisis empírico se realiza utilizando datos de panel de 52 empresas que cotizan en bolsa europea, indexadas en el EUROSTOXX 50, durante el período 2014–2024. Esta muestra transnacional aumenta la relevancia de los resultados al ofrecer una perspectiva más amplia y generalizable que va más allá de los índices bursátiles específicos de cada país.

Los resultados de los modelos de predicción que el comportamiento agregado de búsquedas online posee un poder predictivo significativo, particularmente en cuanto al volumen de negociación y la volatilidad. Además, dicha capacidad es mayor a medida que se amplía el horizonte de predicción. Así, la tesis concluye que las señales tempranas de atención de los inversores, reafirmando el valor que éstas tienen en la predicción financiera y, consecuentemente, en el análisis y en la toma de decisiones de inversión.

Basándose en estos resultados, la tesis valida el GSVI como un proxy confiable de atención inversora en los mercados bursátiles europeos, apoyando la

integración de datos no tradicionales en modelos de finanzas conductuales. En concreto, en lo que respecta al *nowcasting*, si bien su capacidad para captar una percepción inmediata del mercado financiero es limitada, hay que destacar su valor como un indicador complementario que anticipa cambios a muy corto plazo en el retorno, la volatilidad y el volumen de negociación.

Además, la tesis resalta también la importancia a la hora de decidir cómo construir el índice agregado de búsquedas. De esta forma, la tesis aporta conclusiones relevantes cómo distintas alternativas de elaboración del GSVI están alineadas con las variables financieras objeto de estudio y los diferentes horizontes temporales de predicción. Estas conclusiones tienen implicaciones prácticas para la gestión de riesgos de cartera, estrategias de trading basadas en el sentimiento y el desarrollo de señales de mercado más efectivas. Finalmente, la tesis plantea líneas futuras de investigación para mejorar el uso de métricas digitales de atención en los mercados financieros.

Palabras clave: Atención del inversor, finanzas del comportamiento, teoría de la inversión, análisis de datos, Google Search Volume Index (GSVI), mercados bursátiles europeos.

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to Universidad Católica de Murcia (UCAM) for the opportunity to pursue this doctoral thesis. My sincere thanks also go to the University of Stavanger, where I had the privilege of completing a research stay that became a turning point in the development of this thesis. I am especially grateful to Dr. Peter Molnar and Dr. Hamid Cheraghali for their guidance, generosity, and belief in my work.

To my supervisors, Rocío and Nico.

Rocío, thank you for always offering a helping hand whenever I needed it. Your generosity, knowledge, and unconditional support have been invaluable through all these years. You have been a source of calm when I have felt overwhelmed and a source of strength when I have lacked it. Nico, your feedback has pushed this thesis far beyond what I thought it could become. Thank you for your thorough reviews, your insistence on precision, and for making this more than just a thesis. You've also taught me the importance of defending my work, and somehow, you've always known what I needed even when I couldn't articulate it myself. Thank you for recognizing strengths in me that I often fail to see. To both of you, your commitment, patience, trust and dedication have provided me with the tools I needed for this journey. Thank you.

To my father, my mother, and my sister Ana, my favourite person in life. Thank you for always believing in me, for being my safe place, and for reminding me that I can fly while keeping my roots strong. I've reached the end of the staircase! Your love has been the foundation of this entire journey. I would also want to dedicate this thesis to my grandmother Carmen, for being there with her quiet love and presence.

To Alex, thank you for your support and kindness throughout the years. I am truly grateful for the way you stood by me and helped carry this dream forward.

Finally, to my two guardian angels in this journey, Conchi and Nina. This truly would not have been possible without you. Thank you for taking care of me, lifting me up, and being there no matter what. You are women I deeply admire and aspire to be like, and I will forever carry your strength with me.

To my friends, family and colleagues, who has been part of this chapter in one way or another, thank you from the bottom of my heart.

GENERAL TABLE OF CONTENTS

ABSTRACT	7
I - INTRODUCTION.....	29
1.1. Problem Definition	29
1.2. Market Delimitation	32
1.3. Structure Of The Dissertation.....	34
II - STOCK MARKET THEORY	41
2.1. Stock Markets.....	41
2.1.1. Stock market main features: Return, Volatility and Trade Volume.....	43
2.1.1.1. Return.....	44
2.1.1.2. Volatility	45
2.1.1.3. Trading Volume.....	46
2.1.2. Neoclassical Economics and Rational Expectations Theory	47
2.1.3. Efficient Market Hypothesis	49
2.1.4. Behavioural finances and market inefficiency.	53
III - INVESTOR ATTENTION AND GSVI.....	57
3.1. Investor Attention Theory	57
3.2. Measuring Investor Attention	60
3.2.1. Measurements of Investor Attention.....	60
3.2.1.1. Indirect measures	61
3.2.1.2. Direct measures	64
3.2.2. <i>Google Trends</i> and GSVI	67
3.3. Literature Review On Investor Attention And GSVI	70
3.3.1. GSVI Building criteria: Keyword Alternatives and Data Frequency.....	74

3.3.1.1.	GSVI Keyword Alternatives	74
3.3.1.2.	GSVI data frequency	75
3.3.2.	GSVI studies by market region.....	77
3.3.3.	GSVI and financial variables: Return, Volatility and Trading Volume..	78
3.3.3.1.	GSVI and Return.....	79
3.3.3.2.	GSVI and Volatility	81
3.3.3.3.	GSVI and Trading Volume	82
IV -	DEVELOPMENT OF RESEARCH HYPOTHESIS	85
V -	METHODOLOGY.....	91
5.1.	Data Collection	91
5.1.1.	Constituents Data	91
5.1.2.	<i>Google Trends</i> Data	94
5.1.2.1.	Keyword selection.....	94
5.1.2.2.	GSVI.....	95
5.1.3.	Financial Variables	102
5.1.3.1.	Stock Return	103
5.1.3.2.	Volatility	106
5.1.3.3.	Trading volume	108
5.2.	Research Models.....	110
5.2.1.	Panel data.	110
5.2.2.	Cross-sectional and Predictive models.	113
VI -	EMPIRICAL RESULTS	117
6.1.	Summary statistics	117
6.2.	Panel data regressions	119
6.2.1.	GSVI and Return.....	120
6.2.2.	GSVI and Volatility	130
6.2.3.	GSVI and Trading Volume	138

GENERAL TABLE OF CONTENTS		17
VII -	DISCUSSION	151
7.1.	Return	154
7.2.	Volatility	157
7.3.	Trading Volume.....	159
VIII -	CONCLUSIONS, LIMITATIONS AND FUTURE RESEARCH.....	165
8.1.	Conclusions and implications.	165
8.2.	LIMITATION AND FUTURE RESEARCH	171
IX -	BIBLIOGRAPHICAL REFERENCES.....	177
X -	APPENDIXES.....	200

ABBREVIATIONS

AGSVI – Abnormal Google Search Volume Index

AIA – Abnormal Institutional Attention

AMEX – American Stock Exchange

BSE SENSEX – Bombay Stock Exchange Sensitive Index

CAPM – Capital Asset Pricing Model

CESEE – Central, Eastern and South-eastern Europe

EDGAR – Electronic Data Gathering, Analysis, and Retrieval

EMH – Efficient Market Hypothesis

GARCH – Generalized Autoregressive Conditional Heteroskedasticity

GDP – Gross Domestic Product

GSVI – Google Search Volume Index

LSEG – London Stock Exchange Group

OECD – Organisation for Economic Co-operation and Development

PPH – Price Pressure Hypothesis

LIST OF FIGURES

Figure 1. Thesis structure.....	23
Figure 2. Market efficiency degrees forms	50
Figure 3. Salience information process	59
Figure 4. Investor attention measures.....	61
Figure 5. Search intensity of the term Christmas from 2004-2024.	96
Figure 6. Structure of time-retrival of raw Google Trends data.	98
Figure 7. Search volume Raw Data of three companies.....	99
Figure 8. Standarized Google Search volumes using AGSVIDA8, t ; AGSVIDA52, t and AGSVIBIJL52, t	100
Figure 9. Representation of the matching process of GSVI and financial variables.	102
Figure 10. Hausman test.	112

LIST OF TABLES

Table 1. Literature review summary	70
Table 2. Literature review classification based on financial variable, Keyword, Region and Frequency.	73
Table 3. Country distribution of EUROSTOXX50 constituents.	93
Table 4. Sector distribution of EUROSTOXX50 constituents.	93
Table 5. List of variables and description.....	102
Table 6. Classification of Articles by Market and Volatility Approach	107
Table 7. Descriptive statistics for all variables.....	118
Table 8. Correlation Matrix for all the variables.....	119
Table 9. Summary results on AR and ER	120
Table 10. Regression results on Abnormal Return with the AGSVI as independent variable.	122
Table 11. Regression results on Excess Return with the AGSVI as independent variable.	123
Table 12. Regression results on Abnormal Return with the AGSVI and lagged AR as independent variable.	124
Table 13. Regression results on excess return with the AGSVI and lagged excess return as independent variable.....	125
Table 14. Regression results on abnormal return with various independent variables	128
Table 15. Regression results on Excess return with various independent variables.	129
Table 16. Summary results on volatility	130
Table 17. Regression results on Volatility with the AGSVI as independent variable.	132
Table 18. Regression results on volatility with the AGSVI and lagged volatility as independent variable.....	134

Table 19. Regression results on volatility with lagged AR, volatility, ATV, DTV and the AGSVI.	136
Table 20. Regression results on volatility with lagged ER, volatility, ATV, DTV and the AGSVI.	137
Table 21. Summary results on trading volume.	138
Table 22. Regression results on abnormal trading volume with AGSVI as independent variable.....	140
Table 23. Regression results on detrended trading volume with AGSVI as independent variable.....	141
Table 24. Regression results on abnormal trading volume with AGSVI $ATV_t - 1$ as independent variable.....	142
Table 25. Regression results on DTV volume with AGSVI and lagged DTV as independent variable.....	143
Table 26. Regression results on abnormal trading volume with lagged return, volatility, ATV and AGSVI.....	146
Table 27. Regression results on DTV with lagged with lagged return, volatility, DTV and AGSVI.....	147
Table 28. Hypothesis overview.....	152

APPENDIXES

APPENDIX 1. List of companies included in the data sample.	200
---	-----

I – INTRODUCTION

I - INTRODUCTION

1.1. PROBLEM DEFINITION

Stock Markets behaviour is one of the most studied areas in the financial field. Over several years, the research community has study financial asset movements and factors influencing stocks markets (Tajmazinani et al., 2022; Lai et al., 2022; Papadamou et al., 2022; Subramaniam and Chakraborty, 2021; Perlin et al., 2017; Bonfiglioli and Favero, 2005). These movements have been studied from two different perspectives. On the one hand, traditional financial theories focus on rationality and efficiency, assuming that investors are rational and always act to maximize utility (Markowitz, H.M., 1991; Fama, 1970). On the other hand, behavioural finance approaches give room to the psychological factors that influence investors' decisions and, consequently, stock market movements (Hu et al., 2021; Debata et al., 2018; Liu, S., 2015).

Stock market movements serves as an essential barometer of economic vitality and a key tool for informed decision-making at various levels. These movements in financial markets operate on a bidirectional manner. By analyzing these them, investors can make informed decisions, manage risk and optimize returns. Conversely, these movements are driven by other investors' decisions, who react to various data sources available to them.

As it was stated by Vlastakis and Markellos (2012) "Information is the most valuable and highly sought asset in financial markets" (p.1808), emphasizing the importance of information as a valuable resource that influences the outcomes in financial transactions.

The globalization of financial markets facilitated by increased internet usage has enabled rapid information exchange among investors, creating new sources of knowledge. However, although it may seem advantageous, the other side of the coin is that, due to their human nature, investor attention is limited (Shen and Wang, 2023; Chen et al., 2020; Kahneman 1973). Thus, despite investors having access to an extensive amount of information, their cognitive capacity is

constrained, enabling them to process only a portion of it. Consequently, they may overlook subtle but potentially important information (Mondria et al., 2010).

This relation was discussed by the Nobel Prize economist Herbert Alexander Simon (1971), who stated that “a wealth of information creates a poverty of attention and a need to allocate that attention efficiently among the overabundance of information sources that might consume it”. As a result, investors selectively choose financial data and focus only on assets that grab their interest, disregarding fewer appealing options (Da et al., 2010).

Building on this concept, Baber and Odean (2008) formulated the *Price Pressure Hypothesis* (PPH), suggesting that investors tend to be net buyers of stocks that attract attention. They argue that heightened market attention can temporarily boost stock prices, thereby affecting their returns and liquidity positively (Filippou et al., 2023; Chen et al., 2022).

Given that investors have limited attention, financial information is processed in an imperfect manner. This premise stands in contrast to traditional finance theories, which posits that stock prices incorporate all available information as it enters the market, thereby accurately reflecting investors' interests and intentions. In contrast, this thesis is grounded in the principles of behavioural finance, which challenges the notion of fully rational markets. Behavioural finance emphasizes that due to cognitive limitations, investors do not process information optimally, leading to market inefficiencies. These inefficiencies arise from asymmetries in information availability and processing, as well as from predictable biases in investor behaviour.

Based on the preceding assumption, if investor attention is measured, their investment intention can be understood, which will be reflected in financial markets. Therefore, determining investor attention could contribute to explaining stock market variables such as stock returns (Piñeiro et al., 2020; Swamy et al., 2019), volatility (Dimpfl and Jank, 2016; Hamid and Heiden, 2015), trading volume (Kim et al., 2019; Moussa et al., 2017a,b) or liquidity (Aouadi et al., 2017; Ding & Hou, 2015).

To measure investor attention, it is necessary to consider whether the proxy variable represents information demand or supply. In other words, whether it measures what is offered to the investor or what the investor actively seeks. For

example, the publication of specific information about a company does not guarantee that the investor will receive it. However, the active demand of information by investors signals an intention (Da et al., 2011; Joseph et al., 2011).

One of the recent proxy variables to measure information demand is the Google Search Volume Index (GSVI). Google Trends provides search volumes time series of different keywords that captures investors active information demand (Da et al., 2011). The GSVI measures investor interest based on how often specific keywords are searched. The keywords used for building the GSVI are numerous, such as company names, ticker symbol, finance-related term, stock ticker or a combination of the previous mentioned.

The idea is as follows: Google engine is used by investors before trading because they look for stock-related information. Hence, data obtained from Google search queries can be considered as a variable proxy for measuring investor attention (Bank et al., 2011). Some contributions are shown by Cheraghali et al. (2023), Ekinici and Bulut (2021), Kim et al. (2019), and Vozlyublennaya (2014), who use the variable GSVI in forecasting and explanatory stock market models, in relation with other financial variables such as abnormal returns, volatility, and trading volume.

The use of the GSVI is widespread. However, there is ongoing debate about the consistency of the GSVI's effectiveness as a proxy for investor attention in forecasting and nowcasting stock market movements on a global scale. This debate focuses on the findings of research regarding two main ideas: the current and future relationship between the GSVI and financial metrics, and the ability of the GSVI to enhance stock market prediction models. The results obtained have been varied, and the relationship between investor attention and stock market performance has differed based on factors such as the criteria used to construct the index, such as the keywords selected or data frequency, and methodological considerations, including the models applied or the construction of financial variables (Ayala et al., 2024).

Various authors (Da et al., 2011; Joseph et al., 2011; Takeda et al., 2014; Adachi et al., 2017; Swamy et al., 2019; Nguyen et al., 2020; Chen et al., 2022) have conducted research on US and Asian markets, finding that an increase in search volume is positively correlated with an increase in stock returns. In contrast, this

relationship is observed to be negative in investigations by Piñeiro et al. (2020), El Ouadghiri et al. (2018), Perlin et al. (2017), Bijl et al. (2016) and Preis et al. (2013). With respect to volatility and trading volume, the majority of analyses conducted in US, Asian and European markets reveal a strong and significant positive relationship between attention demand and trading volume and volatility. Additionally, numerous studies suggest that the use of the GSVI enhances return forecasts, thereby contributing to profit opportunities associated with search engine data and stock markets (Salisu and Vo, 2021; Lyócsa et al., 2020; Huang, et al., 2019).

The primary research question of this thesis is to determine the current and future relationship between investor attention, approximated through Google search data, and European stock markets. In other words, this thesis wants to test whether a movement in search attention is related to a variation in stock market metrics. To address this question, three pairs of research hypotheses are proposed, which will be developed in subsequent sections, to examine this relationship in greater depth. Each pair of hypotheses focuses on a specific financial variable, analysing the relationship between the GSVI and return, volatility and trading volume. Ultimately, this thesis seeks to contribute by reinforcing the usefulness of internet search data as a powerful tool for both forecasting and nowcasting stock market performance.

1.2. MARKET DELIMITATION

This thesis aims to investigate the capacity of Google searches, as a measure of investor attention, in forecasting and nowcasting financial market movements. Unlike other studies, this thesis is focused on the European Market for the following reasons.

Firstly, the robustness of European market, characterized by companies with high trading volumes and strong economies such as Germany, UK and France, which hold significant prominence on the global stage (Chang et al., 2012). The European economy ranks as the third largest in the world, accounting for 15.2% of the global GDP, following China in the first position (18.9%) and the United States

(15.5%). Additionally, Germany (3.4%), France (2.4%), the UK (2.3%), Italy (1.9%), and Spain (1.4%) collectively contribute 11.4% to the world GDP (World Bank, 2020).

Many of the world's largest multinational companies are listed on European stock markets. Analyzing these markets provides valuable information about the corporate performance and economic prospects of these firms. The companies are widely recognized by the population and more closely followed by retail investors, thereby attracting greater attention from financial media. This will facilitate the sample we obtain from search queries for measuring investor attention, mitigating the lack of data for less-searched keywords (Ramos et al., 2019).

Furthermore, European financial markets are significant recipients and emitters of capital flows. Examining these markets aids in comprehending investment dynamics and international capital movements, which are essential for global investors. Based on the domestic market capitalization, the London Stock Exchange accounts for £4.4 trillion market capitalization (LSEG, 2024); Frankfurt Stock exchange €1.9 trillion (CEIC, December 2023); Euronext €6.9 trillion (Euronext, 2024).

Secondly, there is a lack of academic productivity over European markets on the use of the GSVI. According to the systematic review conducted by Ayala et al. (2024), the majority of studies (41,1% of their sample) are developed in US markets, specifically on the S&P 500 (29%) and DJIA (20%) indices. Given that the US stock market is one of the most actively traded exchanges globally, companies listed on this market illustrate how financial metrics respond to new information. Based on the same study, research conducted on European markets ranks third (after Asia), accounting for 19,6% of their total sample. Notably, only two articles frame the European market under the EUROSTOXX50 stock index. Ramos et al. (2019) measured the effectiveness of using the GSVI on companies belonging to the EURO STOXX 50 index from 2004 to 2011, concluding that an increase in investor attention leads to an increase in volume and volatility. Using the same index, Lobão et al. (2017) obtained weak evidence that current search intensity was related to abnormal stock returns in the following month.

Thus, this study will extend the literature productivity of the performance of the GSVI in the European market.

Lastly, employing a multi-country analysis mitigates potential biases when assessing the impact of the GSVI on financial markets (Perlin et al., 2017). The European region encompasses a diverse and cross-border financial environment, including a variety of financial markets, regulations, banking practices, levels of economic development, and fiscal policies. This heterogeneity can impact investments, trade, and capital flows, as financial conditions and opportunities vary significantly among different European countries. Examining a sample of these markets can offer a more comprehensive understanding into the investor attention theory.

1.3. STRUCTURE OF THE DISSERTATION

This thesis analyses the relation of investor attention, measured using Google search data, and stock market movements. To approximate investor attention, the GSVI is employed, which represents the frequency of a specific keyword searched over a given period. The study is conducted using various estimation methods for both financial variables and the GSVI, aiming to assess the consistency of Google searches in predicting movements in returns, volatility, and trading volume.

The thesis is structured into eight chapters to facilitate comprehension and clearly delineate each section and topic. The structure of the dissertation is presented in Figure 1.

The first chapter aims to clarify the foundations upon which this thesis is based, addressing the problem definition and market delimitation. Additionally, the structure that this dissertation will follow is presented.

Chapter 2 focuses on stock markets as a subject of study, providing a theoretical background about the main financial metrics and their role within stock markets. Furthermore, this chapter also covers various financial theories regarding market efficiency and inefficiency, providing a comparison between traditional asset pricing models and behavioural finance theories. The Efficient Market Hypothesis (EMH) is presented as a cornerstone of market efficiency theories, asserting that markets incorporate all available information, making it impossible to consistently predict asset prices or outperform the market. In contrast,

behavioural finance theories challenge the assumptions of rationality and market perfection emphasizing the irrational behaviours, cognitive biases, and emotional responses of investors, which can lead to systematic deviations from fundamental values and create inefficiencies in the market.

Chapter 3 focuses on the detailed description of the Attention Theory, Limited attention constrains and Investor attention, which together form the foundational assumptions of this thesis. This section also reviews the direct and indirect investor attention measures, ending with the introduction of the GSVI, which approximates investor attention based on Google search queries. This chapter also aims to provide an overview of the general characteristics of the *Google Trends* tool, which is essential for creating the study's variables. Additionally, it delves into the applications of Google search volumes in economics and finances, highlighting the importance of this variable and the outcomes that its data can provide. The chapter underscores the growing relevance of the GSVI as a valuable data source for research in these fields, given its ability to capture real-time shifts in public interest and attention. Lastly, this chapter will review the current literature about the use of the GSVI as a proxy for investor attention and its impact on stock market dynamic.

Chapter 4 develops the research hypothesis that guide the empirical analysis of this thesis. This chapter introduces three pairs of hypotheses, each pair examining the relationship between the market variables (return, volatility and trading volume) and the GSVI. The structure of each pair follows a dual approach: a nowcasting hypothesis, which explores the contemporaneous relationship between the GSVI and the market variable, and a forecasting hypothesis, which tests whether the GSVI has predictive power over the future behaviour of that variable.

Chapter 5 focuses on the methodology and design of the study, beginning with the constituent sample selection. This chapter also covers the keyword selection process, a critical step in accurately capturing relevant data through the GSVI. Detailed instructions are provided on the construction of the GSVI variable and the financial variables necessary for the analysis. Furthermore, the chapter establishes two distinct analytical approaches: the nowcasting approach, which examines the current relation between the variables, and the forecasting approach,

which explores the predictive power of Google searches over future financial outcomes.

Chapter 6 presents the main results of the empirical study, testing the research hypotheses proposed. Results section is structured base on the financial variable, including the findings on the three models proposed.

Similarly, Chapter 7 covers the research discussion, comparing and contrasting this thesis's results with the current literature.

Finally, Chapter 8 presents the conclusions of the dissertation and discuss the main challenges, suggestions for future research, and the main limitations of the study.

CHAPTER I. Introduction

- Problem definition, market delimitation and research contribution.

CHAPTER II. Stock Markets Theory

- Stock Markets
 - Stock market main features
 - Neoclassical economics and Rational Expectation Theory
 - Efficient market hypothesis
 - Behavioural finances and market inefficiency

CHAPTER III. Investor Attention and GSVI

- Investor attention theory
- Measuring investor attention
 - Indirect measures
 - Direct measures
 - *Google Trends* and GSVI
- Literature review on investor attention and GSVI

CHAPTER IV. Development of research hypothesis**CHAPTER V. Methodology**

- | | |
|---|--|
| <ul style="list-style-type: none"> • DATA <ul style="list-style-type: none"> ○ Constituent sample ○ Keyword selection ○ GSVI ○ Financial variables | <ul style="list-style-type: none"> • RESEARCH MODEL <ul style="list-style-type: none"> ○ Panel data ○ Cross-Sectional and Predictive models |
|---|--|

CHAPTER VI. Results**CHAPTER VII. Discussion****CHAPTER VIII. Conclusions, limitations and future lines of research***Figure 1. Thesis structure*

II – STOCK MARKET THEORY

II - STOCK MARKET THEORY

2.1. STOCK MARKETS.

In financial markets, financial assets are exchanged between different economic agents. The primary function of financial markets is to provide liquidity to the economy. On the one hand, through the sale of financial assets, whether equity or fixed income, capital seekers can obtain funds more cheaply than through other external financing options, and suppliers can invest their surplus or savings in a more profitable manner. On the other hand, a liquid market provides investors with a greater ability to exchange financial assets, increasing the trading volume of securities and, consequently, the movements within the markets. (De Mendoza, F., 2008)

The two main features of financial markets are the yield and the risk borne by investors. Given the large number of securities and financial instruments available in financial markets, investors can choose where to invest their capital, taking into account the existing risk-yield relationship of each asset. In this way, investors can efficiently share risk according to their investor profile in relation to their risk tolerance of the investment.

The economic function of the banking system and financial markets has long been understood. In particular, stock markets play a crucial part in this system by offering individuals and businesses valuable resources and opportunities. Stock markets are just as essential and valuable as financial markets, as they provide an alternative where individuals can freely choose among various risk-return options to manage their personal finances. Another advantage of stock markets is their capacity to value companies and assess their projects. When a publicly traded company makes a strategic move, investors can react positively or negatively, with this impact representing the collective sentiment of market participants (Kumari et al., 2024; Cready and Gurun, 2010).

Stock markets have been the subject of extensive academic research due to their fundamental role in the global economy and their inherent complexity (Dhingra et al., 2024; Cohen and Kudryavtsev, 2012; Elton et al., 2004; Jorion and

Goetzmann, 1999; Mitchell and Mulherin, 1994). Researchers from various disciplines, such as economics, finance, and statistics, have sought to understand and model the behaviour of these markets. Some of the most prominent areas of research include market efficiency (Rangapriya and Lokhande, 2023; Varamini and Kalash, 2008; Karemera et al., 1999; Fama and French, 1993) price formation (Bose, S., 2007; Forsythe et al., 1999; Fama, 1970), volatility (Izzeldin et al., 2021; Chaudhary et al., 2020; Vlastakis and Markellos, 2012; Alberg et al., 2008) and the factors that influence investor decisions (Zhao and Zhang, 2024; Swamy et al., 2019; Desagre and D'Hondt, 2019; Da et al., 2011).

Within the areas of study, the question of whether financial markets can be predicted is highly relevant and has attracted considerable research attention. There is an ongoing debate between theories that support market efficiency, and thus the inherent unpredictability of markets, and behavioural finance theories, which argue that psychological factors and cognitive biases play a significant role in market behaviour. The EMH (Fama, 1970) holds that stock prices reflect all available information, making it challenging for investors to achieve consistently superior returns without taking on additional risk. In contrast, behavioural finance challenges this view, emphasizing how factors like investor attention, limited information processing, and irrational decision-making can lead to price anomalies and market inefficiencies. These behavioural perspectives are central to this thesis, as they provide shed light on how limited attention and information constraints affect investor decisions and market dynamics.

Furthermore, the development of quantitative and econometric models has allowed researchers to analyse these phenomena from an empirical perspective. Data analysis methodologies such as time series analysis (Choi et al., 1999; Hjalmarsson, 2010; Kim et al., 2019), GARCH models (Ramos et al., 2019; Herwartz, 2017), and asset pricing models (Takeda and Wakao, 2014; Adachi et al., 2017; Huang, 2019) are widely used to predict volatility, returns, trading volume, and the risks associated with stock market investments.

At the same time, globalization and technological advancements have transformed the nature of financial markets, integrating markets from different countries and facilitating access to information. This has generated new areas of research, such as the impact of international capital flows (Vilutiene and

Dumciuviene, 2020; Evans and Hnatkovska, 2014) and high-frequency trading (Malceniece et al., 2019; Karkowska and Palczewski, 2023; Zhang et al., 2023). However, not all investor behaviour is driven purely by rational decision-making. Increasingly, behavioural finance has explored how emotions and irrational decisions, often influenced by psychological factors, can significantly affect asset prices, challenging traditional market theories.

One of the foundational concepts in stock markets is the risk-yield relationship, which refers to the principle that higher potential returns are typically associated with higher levels of risk. This relationship is key to understanding how investors make decisions and how markets function. The interplay between risk and yield informs nearly every aspect of market behaviour, from individual stock selection to portfolio management.

Investors are generally risk-averse (Sharpe, W., 1965), meaning they require higher returns as compensation for taking on greater risks. In finance, risk is understood as the uncertainty surrounding the returns on an investment, and this uncertainty can be quantified and managed (Ghysels et al., 2005). The greater the potential volatility of an asset's return, the more uncertain the investor is about future earnings. As a result, the higher the potential reward needs to be to justify taking on that risk. Thus, this trade-off shapes decision-making in stock markets, portfolio management, and overall financial strategy (Lundblad, C., 2007; Bali and Peng, 2006; Nyberg, H., 2012).

2.1.1. Stock market main features: Return, Volatility and Trade Volume

In analysing stock markets, three core features provide essential understanding of market behaviour: return, volatility, and trade volume. These elements serve as the foundation for grasping the dynamics of stock prices, investor decision-making, and market sentiment.

Return reflects the gains or losses from an investment, representing the primary goal of most investors. Volatility captures the uncertainty or risk associated with price fluctuations, indicating the stability or instability of an asset's value. Lastly, trade volume measures the liquidity of a stock, revealing the level of market activity and investor interest. Together, these features offer a

comprehensive picture of the stock market, helping investors evaluate opportunities, assess risk, and navigate the complexities of financial markets.

This thesis focuses on measuring the relation between investor attention and stock market movements through three financial variables: return, volatility, and trading volume. In this section, these three variables will be introduced.

2.1.1.1. *Return*

In financial markets, return refers to the gain or loss made on an investment over time. One of the key drivers of return is the fluctuation in asset prices, which can be influenced by various factors. In Andersen (1996) words “price movements are caused primarily by the arrival of new information and the process that incorporates this information into market price”. For example, earnings reports, economic data, or geopolitical events can impact investor sentiment, leading to price adjustments. As this new information is absorbed by the market, it causes shifts in prices, which in turn affects the overall returns on investments.

To gain more precise insights into market performance, researchers often examine Excess Returns and Abnormal Returns, which provide a detailed perspective beyond just basic returns. These metrics help distinguish between normal market fluctuations and significant deviations, offering more reliable information for in-depth market analysis.

- Excess Return is the return above a benchmark, such as the risk-free rate. It allows researchers to measure how much additional reward is gained for taking on higher risk, offering a clearer picture of an asset’s performance relative to safer investments (Cheraghali et al., 2023; Swamy et al., 2019; Bijl et al., 2016).
- Abnormal Return goes a step further by comparing the actual return to the expected return, based on models like the Capital Asset Pricing Model (CAPM). This helps researchers detect whether a stock’s performance aligns with expectations or has been influenced by unexpected events (Tazi et al., 2022; Kim et al., 2019; Adachi et al., 2017).

This thesis proposes the model including both Abnormal and Excess return as financial variables. In this manner, the research is enriched as it explores both

options, providing a more comprehensive analysis of how investor attention impacts stock returns.

2.1.1.2. *Volatility*

Volatility provides essential information about the risk associated with a given asset or market. It allows researchers to identify periods of market turbulence, assess investor sentiment, and quantify uncertainty. Furthermore, understanding volatility helps in evaluating market efficiency and in designing risk management strategies, such as hedging against potential losses.

The central issue in the study of speculative assets is identifying the factors that drive price movements. Are these movements driven by fundamental economic shocks, or do they result from psychological factors such as speculative enthusiasm and shifts in investor sentiment? By measuring volatility, researchers can gain a clearer insight into how investor sentiment and the flow of information influence market instability (Vlasktakis and Markellos, 2012).

This knowledge is crucial not only for predicting periods of heightened risk but also for refining models of price dynamics, enabling the distinction between volatility driven by behavioural factors and that caused by fundamental economic changes.

There exist three main types of volatility measure:

- **Historical Volatility:** This refers to the standard deviation of past asset returns over a specific time frame. It provides insight into how much the asset's price has fluctuated in the past and helps predict future price behaviour based on historical trends. It is a key input in risk management and portfolio optimization models (Ramos et al., 2019; Lahmiri, S., 2017; Aouadi et al., 2013; Brandt and Kinlay, 2005).
- **Implied Volatility:** Derived from options prices, implied volatility reflects the market's expectations for future price movements. Unlike historical volatility, it is forward-looking, providing an estimate of how volatile the market expects the asset to be in the coming period. In other words, it is influenced by investor perceptions of upcoming events. (Basistha et al., 2019; Pati et al., 2018; Dzielinski M, 2012; Juang and Tian, 2005).

- **Realized Volatility:** This is calculated by observing the actual price movements over a given time period, often measured on an intra-day basis. It serves as an empirical measure of how much the price of an asset has fluctuated in real-time (Zhang et al., 2021; Gkillas et al., 2018; Chronopoulos et al., 2018; Hamid and Heiden, 2015).

2.1.1.3. *Trading Volume*

Trading volume is a fundamental variable in stock market analysis, representing the total number of shares or contracts traded in a given period. It serves as an indicator of market activity, reflecting both the liquidity of a stock and the level of investor interest. Unlike price-based metrics such as returns or volatility, which primarily capture changes in asset value, trading volume offers a clearer view of the underlying demand and supply dynamics within the market (Lo and Wang, 2010).

The market activity can be proxied in different ways such as aggregate turnover (Chai et al., 2010; Tkac, P., 1999), intra-day patterns (Heston et al., 2010; Gouriéroux et al., 1999) or trading volume (Desagre and D'Hondt, 2019; Anderson and Dyl, 2005). Among various proxies, trading volume serves as the most direct and transparent measure of trading activity, offering a clearer understanding of market engagement. While price-based metrics focus on value, trading volume emphasizes transaction quantity, making it a more reliable indicator of liquidity and market sentiment. Additionally, it minimizes the potential distortions caused by price volatility, thereby more accurately reflecting genuine investor interest in the assets.

High trading volume typically means a stock can be bought or sold more easily without causing significant price changes (Doyne et al., 2004). This is particularly important for institutional investors managing large positions, as it ensures they can execute trades with minimal market impact. Conversely, low trading volume often indicates a lack of market participation, leading to potential price inefficiencies or increased costs for executing trades (Chordia et al., 2011).

Moreover, many traditional financial models have historically neglected trading volume, assuming homogeneity among investors. That is, that all investors

act similarly. In Karpoff, J (1986) words, “Yet there is active trading in most financial markets, indicating that agents are heterogeneous and operate in a changing environment.” Trading volume reflects the diversity of market participants, each responding to different information, strategies, or needs. This heterogeneity drives continuous trading, highlighting the dynamic nature of financial markets where investors with varying perspectives or objectives trade based on their evolving understanding of the market. (Li et al., 2021).

The treatment of trading volume data is essential for the studies models as raw data alone may be influenced by factors unrelated to market sentiment or investor behaviour, such as seasonal effects or general market growth trends.

One widely used method is the calculation of abnormal volume, which involves comparing the trading volume observed during a particular period with a baseline or average volume derived from historical data under typical market conditions. This normalization process enables researchers to detect deviations or spikes in trading volume that may indicate unusual events (Cheraghali et al., 2023; Lai et al., 2022; Nguyen et al., 2020).

Another technique is the application of detrended trading volume, which removes long-term trends in volume data to highlight short-term fluctuations and isolate immediate responses to specific events. By eliminating the underlying growth or decline trends in trading volume, detrended volume allows analysts to observe variations in activity that are not confounded by broader market dynamics. (Swamy et al., 2017; Bijl et al., 2016).

2.1.2. Neoclassical Economics and Rational Expectations Theory

Foundations of modern finance are based on rational decision making. There exist three key assumptions in the neoclassical theory framework: rational preferences, utility maximization and full information disposal (Ackert and Deaves, 2009).

On the one hand, rational preferences are complete, so the individual can compare all different options and set preferences in one over the other. On the other hand, the option choices are transitive, being the individual's preferences

consistent across the different options. This property ensures that the individual's preferences are coherent and logically consistent.

The second assumption is that individuals tend to maximize the utility in terms of preferences or profits. In other words, they follow an optimizing behaviour in order to maximize their satisfaction from a specific outcome. Sen, A. (1997) defined the action of maximization as "not rejecting an alternative that would be better to have than the one chosen" (p.746), which represents the particular relevance of the act of choosing for individuals.

The third assumption is based on the information used for the decision-making process. Individuals maximize their utility function with perfect and immediate access information and complete knowledge. This assumption is a key element in establishing the foundation of this thesis. The basis of this study is the notion that investors' attention is limited, and therefore, it is not possible to process all available information (Kahneman 1973).

An extension of the neoclassical model is the Rational Expectations Theory, posited by John F. Muth (1961). This theory extends the assumption of rationality to the formation of expectations about the future.

The Rational Expectations Theory posits that individuals make reasoned decisions based on all available information and past experiences. This theory asserts that people use available information efficiently, not ignoring relevant data.

Since individuals base their decisions on the available information and their past experiences, their decisions are typically accurate. When their decisions align with the future outcomes, it reinforces the accuracy of their expectations. Conversely, if their decisions prove to be incorrect, they will adjust their behaviour and expectations based on previous errors.

Considering X to be the studied variables (eg. Inflation rate), hence the Rational Expectations Theory established:

$$X^e = X^{of} \quad (1)$$

Where X^e is the expected value of the inflation rate and X^{of} the optimal forecast of X using all the available information. Therefore, the expectation of the

variables will be the same as the optimal predictions using all the available information (Muth, J., 1961).

The Rational Expectations Hypothesis has been documented across different economic fields, such as inflation rates forecast (Coibion et al., 2018; Coibion and Gorodnichenko, 2015), interest rates (Cieslak, 2018) or stock returns (Bordalo et al., 2019; Greenwood and Shleifer, 2014). In the field of finance, the rational behaviour has been extrapolated to the behaviour of investors and markets outcome, leading to the creation of a parallel theory known as the EMH (Fama, 1970).

2.1.3. Efficient Market Hypothesis

The primary function of capital markets is the allocation of resources between borrowers and lenders. Ideally, these processes would be carried out efficiently as long as the market fully reflects all available information to all participating parties. As Eugene Fama states:

“The primary role of the capital market is allocation of ownership of the economy’s capital stock. In general terms, the ideal is a market in which prices provide accurate signals for resource allocation: that is, a market in which firms can make production- investment decisions, and investors can choose among the securities that represent ownership of firms’ activities under the assumption that security prices at any time “fully reflect” all available information. A market in which prices always “fully reflect” available information is called “efficient.” (Fama 1970, p.383)

The EMH assumes that security market prices reflect all available information, so they represent good assessment of the assets and allow a proper allocation of the capital in markets. In order to explain the sense of “all available information”, Fama (1970) proposed three form of market efficiency (see Figure 1). The weak form assumes that security prices already include information about past prices. The semi-strong form suggests that all the information that is public (companies’ financial statement data, stock splits, etc.) is reflected in today’s security prices. Lastly, strong form believes that all public and private information is presented in stock prices, which means that there are no profit opportunities associated to forecasting strategies.

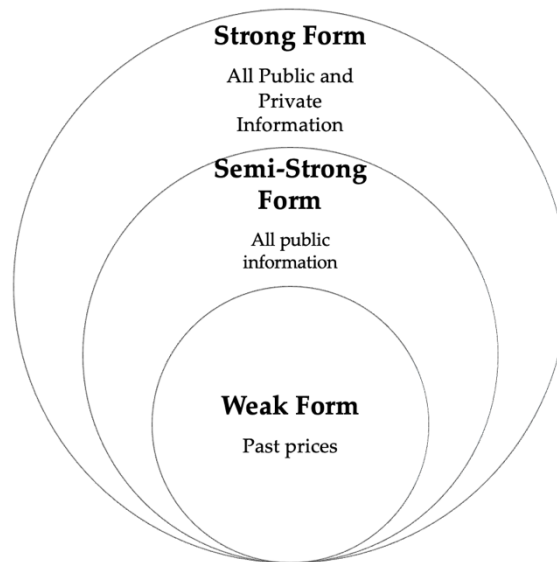


Figure 2. Market efficiency degrees forms

Source: Author's elaboration based on Fama (1970).

An important consideration regarding the EMH is the cost of obtaining information. If the market is efficient because it reflects all available information, then it must be assumed that the cost of that information (trading cost, information cost) is zero. Jensen (1978) explained in his study on the EMH that markets would remain efficient as long as the marginal benefit of acting on information does not exceed the marginal cost of obtaining that information. Hence, the profit made from beating the market cannot be higher than the cost of acquiring the means to do so.

In an efficient market, the stock price of a security is established by the present value of the expected future dividends (see equation 2).

$$P_{it} = \sum_{k=0}^{\infty} \frac{E_t D_{t+k}}{(1+r)^{k+1}} \quad (2)$$

Where P_{it} is the price of stock i at time t , $E_t D_{t+k}$ is the expected value of the future dividend¹ in the time period and r is the discount or risk rate.

In accordance with the Rational Expectations Theory (Muth 1961), all available information about prices is used. If investors receive new information that influences their expectations about a company's future profitability, the stock price will reflect this change almost instantaneously. Hence, in a market where investors have rational expectations, asset prices primarily respond to new information that influences the underlying economic fundamentals factors.

At this point, it is crucial to explain the financial connotation of “receiving information.” As it was suggested in Ackert and Deaves (2009), “for most people, information means knowledge, but in finance when we say ‘information’ we mean items that are truly unanticipated” (p. 30). Consequently, only unexpected information constitutes genuinely new information. If this new information cannot be predicted and stock prices absorb it quickly (because it is an efficient market), then prices, by definition, will be equally unpredictable.

It is thus considered that stock prices in the market follow a “random walk” making it impossible to forecast future market movement (Fama, 1995). This suggests that investments follow a “fair game” paradigm, wherein no investment strategy can systematically outperform the market on a risk-adjusted basis, except by chance.

It is important to emphasize that the EMH does not suggest that stock prices are random; rather, it posits that prices are set accurately based on all current and available information. This means that at any given moment, the price of a stock is a true reflection of its intrinsic value, as the market has already processed and

¹ Assuming that not all stocks pay dividends, the model considers as future dividends all cash flows to shareholders. (Ackert and Smith, 1993).

incorporated all relevant information. What it truly suggests is the inability to predict opportunities for excess returns on future investments.

The EMH has been a challenge for the financial scientific community, as over the years various violations of market efficiency have been highlighted. As Fama (1970) himself notes, “A frictionless market in which all information is freely available and investors agree on its implications is, of course, not descriptive of markets met in practice” (p. 387).

There are several anomalies that question this theory, suggesting that asset prices can deviate from their fundamental values due to factors not directly tied to rational or efficient information processing. These anomalies indicate that markets are sometimes driven by forces other than pure data or logic, causing temporary distortions in asset prices that do not align with underlying value.

One of these anomalies that contradicts the EMH is the small firm effect (Jacobsen et al., 2005; Banz, 1981), which refers to the tendency of smaller companies to generate higher risk-adjusted returns compared to larger firms. According to the EMH, any excess returns should be quickly arbitrated away as all investors would access and act upon the same information. However, smaller firms often receive less attention from analysts and media, leading to limited coverage and, consequently, a slower incorporation of information into their stock prices.

Another anomaly is the information asymmetry (Goel et al., 2021; Alao, 2018, Lambert et al., 2012) that occurs when some investors have access to information that others lack, resulting in asset prices that fail to fully reflect all relevant information. While the EMH assumes that information is freely and equally accessible to all investors, the reality is that obtaining and processing information comes with significant costs, leading to an uneven distribution of information among investors (Dávila and Parlato, 2021; Gu and Hitt, 2021; Elton et al., 1993). Additionally, information asymmetry can also arise from exclusive access to privileged or insider knowledge. These disparities in information access led to price deviations from fundamental values.

The EMH is based on the assumption that investors behave rationally, processing all available information logically to maximize their returns. This research is based on the idea that investor attention is limited (Kahneman, 1973), and therefore, even with access to all available information, they are unable to

process it all. Even if investors have access to all relevant information, their ability to process it is finite, meaning they cannot absorb all the information available in the market. Instead, they focus only on a fraction of it, specifically the part that captures their attention most significantly. This fact undermined the EMH and raises the question of whether and to what extent market prices reflect investors interest. (Huang et al., 2020; Takeda and Wakao, 2014).

This thesis measures investor attention through Google searches and tests how these searches can predict market movements. Using search volumes as a proxy for attention, this approach allows for the observation of how investors react to specific events and how their attention influences asset demand and price changes. Investigations developed by Akarsu and Söer (2022); Nguyen et al. (2020); Pereira et al. (2018) or Da et al. (2011), among others, agree on the predictive capacity of the GSVI. The ability to predict asset prices based on search patterns suggests that markets may be influenced by psychological and cognitive factors that the EMH overlooks.

2.1.4. Behavioural finances and market inefficiency.

Some people think that behavioural finance introduced psychology into finance, but psychology was never out of finance. (Statman 1999, p. 19).

Behavioural finances were born in contrast of traditional financial theories that support investor rationality and market efficiency (Shefrin, 2002; Statman, 1995; De Bondt and Thaler, 1986). According to Konstantinidis et al., (2012):

The inherent weaknesses of the Hypothesis have become the investing weapons of the new scientific approach. Investing rationality and efficient market processes over time contradict investors' psychology, biased behavioural rules and market bubbles. Information efficiency and the reconciling approach of arbitrage have conflicted the inefficient access to investment information and long-term market anomalies (Konstantinidis et al., 2012, p.16).

Although the first theories on behavioural finance started to emerge in the 1980s (Kahneman and Tversky, 1979), it was not until the early 1990s that this field began to gain prominence within the scientific community, not as a complementary

approach but as a confrontational theory to standard finances. In Statman (1999) words, "people are "rational" in standard finance; they are "normal" in behavioural finance" (p.26).

In Wäneryd (2001) terms, behavioural finance is referred "when psychological thinking is applied to financial markets in search of explanations for behaviour that apparently deviates from what it is expected on the basis of economic theory, and efficient-market theory in particular" (p.7). Furthermore, he suggested that it is an approach for "imperfect markets and imperfect people" (p.7). For Thaler (2016), it is considered as "open-minded finance" (p.1597).

Market anomalies and investors' cognitive biases have led the scientific community to question the foundations upon which traditional efficient market theories are built. In Statman (1999) words, "today's standard finance is so weighted down with anomalies that reconstructing financial theory along behavioural lines makes much sense" (p.19).

One of the leading theories in the field of behavioural finance is the Prospect Theory, proposed by Kahneman and Tversky (1979) in their work *Prospect Theory: An Analysis of Decision Under Risk*. The authors suggested that people do not always make rational financial decisions due to cognitive and emotional biases. Furthermore, they analysed how this fact affects decision-making in situations where individuals must choose between alternatives that involve risk.

The Prospect Theory helps explain the irrational behaviour of investors based on the following assumptions:

(1) People tend to overweight small probabilities and underweight large probabilities. In this sense, investors might take on excessive risk with speculative investments, overlook stable and reliable options, engage in excessive trading, and follow herd behaviour, all due to the misweighting of probabilities.

(2) The pain of losing is psychologically about twice as powerful as the pleasure of gaining. According to Kahneman and Tversky (1979), losses loom larger than gains, so investors tend to be more risk-averse for losses than for gains.

(3) The way a decision or problem is framed significantly impacts choices. Framing can affect how information is processed and perceived by investors. If news is framed in a way that emphasizes certain aspects over others, it can lead to

skewed interpretations and temporary inefficiencies, where prices may not immediately reflect all aspects of the new information, contradicting the EMH.

Unlike the EMH, behavioural finance holds that investors process information imperfectly. On the one hand, investors may be influenced by how information is presented, known as framing (Chen and Lai, 2013; Glaser et al., 2007; Levin et al., 1998) or rely on specific reference points, anchoring bias (Owusu and Laryea, 2023; Andersen, 2010; Campbell and Sharpe, 2009), leading to systematic errors in decision-making and pricing inefficiencies. On the other hand, investors' ability to reason information is limited (Kahneman, 1973), making them unable to assimilate all the content and make decisions based on everything presented to them.

In a world where information changes rapidly and markets are globalized, it is challenging to discuss information that is both complete and easily assimilated. This capacity for assimilation is impeded by information overload, which can result in confusion and an inability to reason effectively. According to Reber and Schwarz (1999), information that is easier to process is more likely to be perceived as true.

The incapacity of investors to reason through information is a topic that will be explored deeply in the next chapter.

III– INVESTOR ATTENTION AND GSVI

III - INVESTOR ATTENTION AND GSVI

This chapter will outline the theoretical foundations of investor attention and provide a framework for understanding how attention influences investor behaviour. Additionally, a section will be included to present the main measures of investor attention found in the literature, with a particular focus on the direct measure through Google searches, which is the main variable used in this thesis study.

3.1. INVESTOR ATTENTION THEORY

Attention, as a psychological concept, refers to the cognitive process of selectively concentrating on specific information while ignoring other stimuli (Kahneman, 1973). It is a limited resource, meaning that individuals cannot process all incoming information simultaneously and must allocate their attention based on importance or relevance. This scarcity of attention is a foundational idea in cognitive psychology and was explored extensively by Daniel Kahneman (1973) in his book *Attention and Effort*. Kahneman (1973) proposed that attention functions like a finite pool of cognitive resources, and individuals must manage these resources when performing tasks or processing information. According to the author, attention requires effort, and people instinctively balance their mental resources by focusing on certain activities while ignoring others. This balancing act is what helps maintain efficiency in decision-making but also introduces limitations.

In this attention capacity model, attention is like a reservoir with a limited supply. When a task requires significant focus or effort, more of that reservoir is devoted to it, leaving less available for other tasks. This need to balance attention among competing activities is often referred to as "selective attention." The consequence is, as Shen and Wang (2023) noted, that "attention to one task is a substitution for attention to other tasks" (p.122). It is important to emphasize that this conflict of limited attention does not focus on daily situations such as talking

on the phone while walking home, but rather on central cognitive processes such as memory retrieval and action planning for managing multiple tasks (Peng and Xiong, 2006).

Multitasking forces the limited attention to split between tasks, often resulting in reduced efficiency and an increased likelihood of errors. Research in cognitive psychology has shown that when people attempt to do multiple tasks simultaneously, they may experience switching costs, or the mental effort and time required to shift focus from one task to another (Cardoso-Leite et al., 2015; Wylie and Allport, 2000). This diminishes performance and illustrates the limitations of the attentional capacity.

In his book, Kahneman (1973) emphasized the link between attention and effort, suggesting that sustaining attention on demanding tasks is mentally tiring. This connection is significant because it suggests that attention is not a constant resource as it fluctuates based on individual capacity, the nature of the task, and the individual's mental state.

The concept of limited attention has implications for daily lives, especially in a world saturated with information and demands on our time. Because attention is a finite resource, people often prioritize certain activities over others, consciously or unconsciously. This prioritization means that many stimuli, even those that may be important, are disregarded if they do not capture the immediate focus. For example, while reading an article, a person may miss relevant details due to simultaneous distractions, like checking their phone or thinking about an unrelated task.

In the realm of finance, attention plays a crucial role in how individuals make decisions. Financial markets are filled with overwhelming amounts of information, including market news, stock prices, economic reports, and expert analyses. Because of the finite nature of attention, investors often cannot process all available information, leading them to focus on specific, salient information while neglecting others (Da et al., 2011). Attention constraints also explain certain behavioural finance phenomena, such as herding behaviour, where investors follow popular trends without fully analysing underlying fundamentals. In the research carried out by Devenow and Welch (1996), the authors found that investors that follow other investors feel more trading secure.

When the concept of limited attention is applied to specific behaviours in financial markets, it is referred to as “limited investor attention”. Limited investor attention refers to the idea that investors, constrained by cognitive resources, can only attend to a subset of all relevant information (Da et al., 2011).

Studies that address the inability of investor to reason and process all information due to attention limitations can be found in Peng and Xiong (2006), Sims (2003) and Merton (1987). Peng and Xiong (2006) evidenced the effect of investor’s attention allocation into asset-price dynamic. According to the authors, investors tend to a category-learning behaviour, and they will process more market than firm-specific information. Sims (2003) introduced the model of “rational inattention”, explaining how individuals selectively process information due to cognitive constraints, optimizing their decisions by focusing only on the most relevant data while ignoring less critical information. In the Finance field, Merton (1978) demonstrated that limited attention leads to imperfect portfolio diversification, and as a result, lesser-known firms need to offer higher returns to compensate for the reduced diversification in equilibrium.

Closely related to investor attention theories is the concept of salience of information. In the words of Bordalo et al. (2022), “a stimulus is salient when it attracts the decision maker’s attention ‘bottom-up,’ automatically and involuntarily” (p.1). Based on the idea that there is a daily overload of information, salient information is able to capture more attention. Figure 3 presents the mechanism of how salience of information affects the final decision process.

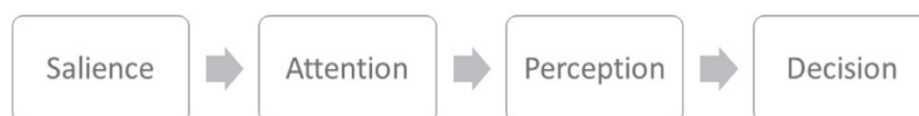


Figure 3. Salience information process

Source: Ramos et al., (2020)

Numerous studies have emphasized the features of salient information that capture investor attention and how this impacts their decision-making. For instance, Ramos et al. (2020) developed a model to examine the salience of both

firm-specific and market-wide information in European markets, aiming to better understand the salience information process. They measured information salience using a high-low indicator and utilized Google search volumes as a proxy for investor attention. Similarly, Hirshleifer et al. (2009) investigated how attention-grabbing events can lead to overreactions or underreactions in stock prices. In the same vein, Barber and Odean (2008) analysed how highly visible information, such as extensive media coverage or significant price changes, can result in biased investor decisions, particularly influencing buying behaviour. Additionally, Bordalo et al. (2012) explained that investors tend to react more strongly to salient information, even when it is not the most relevant for optimal decision-making.

In a purely financial context, and considering that investors' attention is limited, they must be selective in processing the vast amount of information available for investment decision-making. This idea contrasts with traditional models, which assumes that all available information is absorbed and reflected in prices. However, the difficulty in assessing information is exacerbated by the abundance of available data, as ordinary individuals possess inherent imperfections and have limited capacity to process information requirements effectively (Ackert and Deaves, 2009).

Therefore, the concept of investor attention plays a crucial role in financial decision-making, investment choices, portfolio diversification, and the development of investment strategies. The key question at this point is: how can attention be measured? The following section introduces the primary methods used to measure attention.

3.2. MEASURING INVESTOR ATTENTION

3.2.1. Measurements of Investor Attention

Measuring investor attention is a challenge for the research community. The measures are classified into direct and indirect proxies for investor attention, depending on how the information is presented and how the investor receives it. Direct measures typically involve metrics such as search engine queries (Google and Baidu) or social media interactions (X [formerly Twitter]), which provide clear indicators of investor interest in specific assets or market events. In contrast,

indirect proxies might include abnormal trading volume or stock price movements, which reflect investor attention indirectly by signalling how investors react to information. Understanding these distinctions is crucial for accurately assessing the impact of investor attention on financial decision-making and market behaviour.

This section will review the existing literature about how to measure investor attention. Figure 4 presents the structure followed.

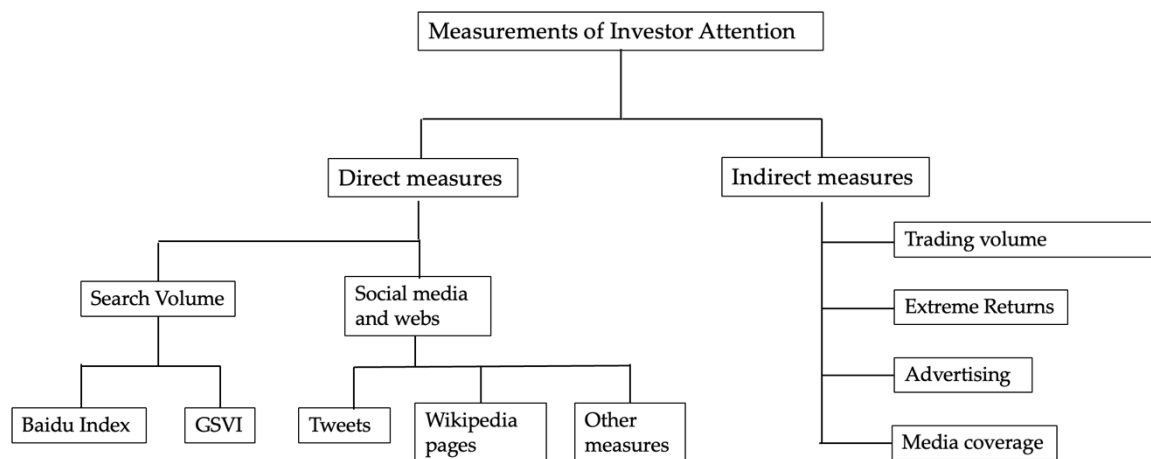


Figure 4. Investor attention measures

Source: Own Elaboration based on Shen and Wang (2023)

3.2.1.1. *Indirect measures*

Indirect measures for investor attention are related to attention-grabbing events such as abnormal trading activity or the release of big firm announcements. These events are defined as “shocks which would attract investor attention within the market” (Shen and Wang, 2023, p. 123). Assuming that investor attention is limited, they are likely to prioritize stocks that stand out due to their distinction

relative to others, as these are more likely to capture their focus based on the salience of the available information.

Indirect measures of investor attention refer to indicators that catch external signals of interest. These measures are based on observable market behaviours that are assumed to reflect underlying investor attention.

Barber and Odean (2008) investigated investor attention using three distinct indirect measures: abnormal trading volume, media coverage, and extreme returns. Through their study of assets listed on the NYSE, AMEX, and NASDAQ, they determined that abnormal trading volume served as a more effective proxy for investor attention compared to media coverage and extreme returns. Other studies that proxied investor attention through trading volume can be found in Lin et al. (2014) and Hou et al. (2009).

Extreme returns were also used as a measure for investor attention. This is based on the idea that large price changes are highly noticeable and more likely to grab the attention of investors (Barber and Odean, 2008; Seasholes and Wu, 2007). However, although these big movements stand out from normal market activity, there is no certainty that these events are capturing investor interest.

Another indirect measure of investor attention is advertising expenditure. Research by Chemmanur and Yan (2009) and Lou (2014) provides evidence that increasing advertising investments can enhance the likelihood of attracting investment in a company's stock, extending beyond merely drawing customers or promoting products. Similarly, scholars have explored the relationship between investor attention and media coverage. An increase on visibility in news and media could increase the likelihood that investors will pay attention to the stock. Some evidence can be found in Koester et al., (2016), Yuan (2015) and Barber and Odean (2008).

With the rise of the Internet and digital technologies, researchers now have access to a wealth of data that offers new ways to directly measure investor attention. Before the availability of these direct data sources, researchers had to rely on indirect measures to approximate investor attention. While useful at the time²,

² Note that the majority of the literature of indirect measures for investor attention is prior to 2016.

these indirect proxies present some intrinsic limitations and often led to imprecise conclusions.

Indirect measures are highly ambiguous and often lack a direct causal link to investor attention. High trading volume, for example, may reflect increased attention, but it could equally indicate other dynamics, such as institutional rebalancing, calendar anomalies or market speculation (Batrinca et al., 2018). Similarly, spikes in media coverage may be driven by sensationalized stories rather than actual investor interest. Because these metrics capture various types of market activity, they blur the line between attention-driven behaviour and other factors, leading to potentially misleading interpretations. This lack of specificity undermines their validity as indicators of investor attention.

Furthermore, indirect measures are susceptible to context-dependent biases. When testing investor attention by media coverage, it can indicate different effects dependent on the asset or market in question. For instance, media coverage of a well-known technology stock could attract significant investor attention because of its popularity, while the same coverage of a lesser-known stock may go unnoticed. The impact of media coverage varies based on the stock's familiarity and notoriety.

Another limitation of indirect measures is that they cannot accurately determine if a person is truly interested, because there is no attention demand being measured (Fink and Johann, 2014; Da et al., 2011). Instead, these measures focus solely on attention to the information supply, meaning they track the amount of information available or how much attention that information receives, but they do not capture whether people are interested in the content. In other words, while you can measure how many people viewed an advertisement, you cannot know if those people were actually interested or if the information truly captured their attention. For example, someone might see a new about a financial asset without being interested in investing in it. The measurement of attention is limited to what was available or exposed to people, but not their actual interest or intent.

The following section will explore the direct measures of investor attention, which offer a more accurate and reliable approach to assessing investor interest and behaviour by capturing explicit actions and intentions.

3.2.1.2. *Direct measures*

Recent technological advancements and the growing popularity of social media have enabled the development of direct methods to assess investor attention. Unlike indirect proxies, these methods, such as tracking online search volume, social media mentions and website traffic, capture the direct demand for information from investors. Such demand can reflect investor interest in particular assets, indicating what investors are focusing on in real-time or over specific periods. This section will review the literature and present the various direct approaches for measuring investor attention.

Direct measures of investor attention include tracking social media activity such as views or interactions. Wikipedia page views serve as direct measures of investor attention. Gómez-Martínez et al., (2022) analyzed the explanatory and predictive capacity of Wikipedia page views for the NASDAQ Index, concluding that the volume of queries related to NASDAQ companies is a statistically significant variable in reflecting the index's evolution. In a similar vein, Pyun (2024) discovered that fluctuations in the Wikipedia page views of firms are valuable indicators of stock performance across all stocks listed on major US stock exchanges. In contrast, Osabuohien-Irabor (2021) investigated the relationship between changes in Wikipedia page views and stock returns, trading volume, and volatility dynamics among companies listed on the Nigerian Stock Exchange. This study found that investor attention does not significantly impact any of the financial metrics examined.

Alternative measures of investor attention in the realm of social media include quantifying the social power via the X platform. Li et al., (2016) proposed asset-specific tweets as a measure of investor attention, highlighting their significance in understanding market behaviour. Building on this, Ganesh and Iyer (2023) analysed the relationship between corporate tweets and stock returns, drawing on data from over 1.2 million firm tweets made by listed companies in the DJIA between April 2013 and July 2020. Their findings suggest that a shift in corporate tweeting behaviour leads to a positive impact on changes in trading volume and stock returns. Suardi et al., (2022) investigated the predictive power of information contained in social media tweets, specifically focusing on bitcoin market dynamics.

Recently, new sources of direct measures of investor attention have emerged, drawing from websites and social media. For example, Iliev et al. (2021) examined attention toward corporate governance by analysing investor views filed in the EDGAR3 online database. From an institutional perspective, Ben-Raphael et al., (2017) introduced the Abnormal Institutional Attention (AIA) variable to compare the responses of retail and institutional investors to attention-grabbing events and their effects on market dynamics.

Beyond social media activity, another valuable approach for measuring investor attention is search queries volumes. When investors actively seek information about a specific stock, such as by researching recent news, looking for financial data, or tracking price changes, they are directing their attention to that stock (Joseph et al., 2011; Ramos et al., 2019). This action of searching for information indicates an immediate interest about the performance of the stock, suggesting that the investor is focusing their resources on understanding or evaluating it. This concept, as discussed by Da et al. (2011), frames investor attention as a form of “information demand”. By actively looking up details about a stock, investors are effectively signalling their interest or intent, whether it is due to a market event or a notable price movement. In this sense, the investor is an active seeker of information demand, constantly searching for signals that might guide their investment decisions.

Measuring this demand for information, therefore, provides a direct indication of which stocks are currently capturing investor focus and may potentially anticipate trading activity (Kim et al., 2019; Bijl et al., 2016; Preis et al., 2013). An increase in searches often translates into greater popularity for a stock, as it signals growing investor interest and attention (Yoshinaga and Rocco, 2020). This surge in search activity can provide early indications of future investor actions, allowing for the anticipation of market movements and potential shifts in financial markets before they fully materialize.

In the current literature, the most commonly used search index to approximate investor attention is derived from search activity on Google. Google is the most popular search engine globally (StatCounter, 2024), leading to a greater

³ Electronic Data Gathering, Analysis and Retrieval system.

number of studies that utilize the GSVI as a proxy for investor attention (Lai et al., 2022; Desagre and D'Hondt, 2019; Chen and Lo, 2019; Dimpfl and Jank, 2016). What makes Google searches particularly powerful in this context is that they provide a genuine reflection of investor attention in real-time. According to Da et al. (2011), the information obtained from the GSVI quickly influences market prices, while the news about a company affects prices more gradually.

Investors may not always express their concerns explicitly in news reports, social media, or financial discussions, but their search queries reveal what is on their minds, offering a direct measure of what they are actively thinking or investigating. These searches can reflect underlying intuitions and emotional responses that influence investment decisions, providing a richer, more nuanced understanding of the factors driving market behaviour. Thus, Google search data could detect potential signals for stock market movements (Huang, 2019).

Just as Google search data reveals interests of people, Baidu Index searches offer a window into the minds of Chinese population. However, although there exists extensive research about the use of Baidu index (Gao et al., 2023; Yang et al., 2021; Zhang et al., 2013), its primary challenge and limitation, compared to Google, is that it is restricted to China and the Chinese internet ecosystem. This creates a narrower, more region-specific sample, as it only captures the behaviour of Chinese users, excluding valuable data from other markets. Contrary, Google offers aggregated search data from around the world, allowing a comprehensive view of investor sentiment across different regions, cultures, and economic contexts, further enhancing the ability to understand global market dynamics and how investors from diverse backgrounds react to the same events (Ahundjanov et al., 2020; Lyócsa et al., 2020; Škrinjarić, 2019).

The use of investor attention, as proxied through Google search volumes, has been tested in relation to various financial variables, including volatility (Smales, 2021; Ramos et al., 2019; Vlastakis and Markellos, 2012), return (Nguyen et al., 2020; Kim et al., 2019; Adachi et al., 2017), liquidity (Ayaz et al., 2021; Xiao and Wang, 2021), and trading volume (Perlin et al., 2017; Moussa et al., 2017a; Tantaopas et al., 2016). Sections 3.2.2 and 3.3 will further elaborate on this measure, which serves as the basis for assessing investor attention in this thesis.

3.2.2. Google Trends and GSVI

Google Trends is a service that offers data series of search volumes for a specific term or word. This tool is of free use and belongs to the Google search engine. Google dominates the search engine market with 82% market share worldwide, followed by *Bing* (10,5%), *Yahoo!* (2,7%), *YANDEX* (2,42%), *DuckDuckGo* (0,7%) and *Baidu* (0,5%) (StatCounter, 2024). However, there are some specific regions that use other leading alternatives as the users of Russia with *Yandex* (63%) (Yandex, 2024) or the Chinese population with *Baidu* (53%) (StatCounter, 2024). Although search volumes are available from 2004, most studies have been conducted with data after 2008 (Challet and Ayed, 2013).

Google Trends also allows to compare the popularity of multiple search terms side-by-side to see which one is more searched for at any given time. Among the key features and functionalities of the engine, it allows users to have a view of trending searches in real-time, showcasing the most popular search queries at a specific moment.

Google search data serves as a highly powerful and valuable resource for the scientific community. The analysis of Google search traffic is applied across various fields of knowledge, including healthcare, macroeconomic trends, sociology, economics and finance. This data provides real-time information about human behaviour, societal trends, and economic activities, making it an indispensable tool for researchers.

The pioneer on applying Google search data was Ginsberg et al., (2009), who used data collected from Google to track influenza-like illness in the US population. According to the authors, Google search queries could improve early detection, which reduce the impact in both seasonal and pandemic influenza. There exist manifold applications of Google search queries. In health, it has been used for monitoring toxicological outbreaks (Yin and Ho, 2012), mental health diseases (Ayers et al., 2013; Yang et al., 2010), tracking infectious disease (Mytton et al., 2012; Seifter et al., 2010) or cancer-related topics (Hopkins and Secrest, 2019; Bloom et al., 2015).

In Economics, Choi and Varian (2012) introduced an innovative approach by leveraging Google search traffic for "*nowcasting*" economic indicators, including employment, motor vehicle sales, travel, and consumer confidence. The authors

define the concept of nowcasting as a contemporaneous forecasting, claiming that Google searches could help in predicting the present. They established foundational benchmarks for future research and provided an introduction to Google Trends for readers.

Google search data has proven useful for predicting and anticipating economic indicators such as GDP growth or unemployment rates. In Petropoulos et al., (2021) words "Social data is more responsive compared to official financial indicator, which are usually available with a lag of several weeks or months" (p.353). Some examples can be shown in Götz and Knetsch (2019), who incorporated Google search data into a bridge equation model for the German economy to assess its impact on GDP growth forecasts. Their findings indicate that including Google searches data led to more accurate GDP growth predictions for Germany. Similar results were found by Ferrara and Simoni (2019), concluding that integrating Google searches improved forecasting accuracy for the Eurozone, even when controlling for official variables. Expanding this application beyond Europe, Bantis et al. (2023) used Google Trends data for forecasting in the US and Brazilian markets, demonstrating that models incorporating Google data performed better than those excluding it.

D'Amuri and Marcucci (2017) tested Google search volumes for predicting the monthly US unemployment rate, concluding that Google-based model outperform other alternatives such as those based on survey's, labour force flows or non-linear models. Comparable findings were revealed by Borup & Schütte (2022), who demonstrated the strong predictability capacity of Google search data for unemployment growth in the US. The data provided by Google Trend is also tested to forecast other specific macroeconomic series such as private consumption (Vosen & Schmidt, 2011) or price levels (Seabold & Coppola, 2015).

Recently, the OECD has presented the OECD Weekly Tracker, a tool that leverages Google Trends search data to monitor economic activity across 46 OECD and G20 countries in near real-time. (Woloszko, N., 2020). The analysis identifies the predictive power of certain search topics, such as "bankruptcies," "economic crisis", "investment", "luggage", and "mortgage", which are particularly relevant to economic conditions. Thus, this Google search-based tracker performs well in capturing economic trends, including during the COVID-19 crisis.

More recent applications of Google searches can be found in Gonzalez-Gallego et al. (2024), who develop a forecasting model based on Google search data for predicting violence against women.

In the financial field, the use of Google search queries can be traced back to the work of Preis et al. (2010), who explored the relationship between search volume data, measured by the names of constituents in the S&P 500 stock market index, and transaction volumes. Building on this foundation, Da et al., (2011) introduced the GSVI as a proxy for measuring investor attention. The GSVI captures the frequency of specific keywords searched on Google, providing valuable insights into public interest and potential shifts in consumer behaviour or sentiment. Since 2011, research on the connection between the GSVI and financial market performance has expanded rapidly, highlighting the growing role of search data in analysing market dynamics.

The applications of the GSVI in finance include stocks, cryptocurrencies or fixed income instruments such as bonds. Zhang et al. (2021) found empirical evidence of the relationship between internet attention, measured by Google search volume, and Bitcoin trading volume. Koch and Dimpfl (2023) conducted an analysis of Google search trends for terms such as "Bitcoin," "Ethereum", "Litecoin" and "Monero," concluding that Bitcoin's search attention is the primary factor influencing cryptocurrency prices. Pham and Huynh (2020) studied the role of investor attention in guiding sustainable investment flows, demonstrating that investor attention significantly influences green bond market performance.

While these applications demonstrate the versatility of the GSVI as a forecasting tool, this thesis remains focused on its predictive and nowcasting power within stock indices, where it continues to play a pivotal role in understanding market dynamics and investor decision-making processes. Section 3.3 will extensively review the current literature about the use of GSVI in analysing stock market performance.

3.3. LITERATURE REVIEW ON INVESTOR ATTENTION AND GSVI

As it has been discussed in previous sections, the GSVI is used in many financial and economics areas. This section will review the current literature about the use of the GSVI as a variable proxy for investor attention and its relationship with three financial variables: return, volatility and trading volume. This literature review includes studies developed over stock markets, not considering papers that uses as main input other financial assets, such as cryptocurrencies, fix income assets or commodities, among others.

The review is carried out in three steps. First, it examines the criteria for constructing the GSVI, including keyword alternatives and data frequency. Second, it includes a section analysing the literature based on the geographical focus of the research. Finally, it discusses literature exploring the relationship between the GSVI and three financial metrics: return, volatility, and trading volume.

Table 1 summarizes 62 papers based on the region of study, stock market index, financial variables, keywords searched and data frequency. Table 2 presents the papers classification based on financial metric and data sample criteria.

Table 1. Literature review summary

	Reference	Region	Market	Financial Variable	Data Frequency	GSVI Keyword	Main Approach
1	Ghaddab et al. (2024)	CESEE region	Hungary, Croatia, Ukraine, Slovenia, Poland, Bosnia, Czech Republic, Slovakia, Bulgaria and Serbia stock index	Return, Volatility	Monthly	INDEX	Nowcasting
2	Sukamulja and Thedora (2023)	China, India, US, Indonesia and Brazil	SSE Composite Index, BSE SENSEX, NYSE Composite, IDX Composite, Brazil Stock Exchange Index.	R, Vol, TV	Weekly	INDEX	Forecasting
3	Cheraghali et al. (2023)	US	S&P Global Luxury Index	Return	Weekly	TICKER	Nowcasting /forecasting
4	Chundakkadan & Nedumparambil (2022)	Multiple	59 Countries index	Return, Volatility	Daily	TERM	Nowcasting
5	Dharani et al. (2022)	India	S&P BSE 500 index	Return	Weekly	TICKER	Forecasting

6	Saxena and Chakraborty (2022)	India	NSE100 index	Return	Weekly	TICKER	Forecasting
7	Ouadghiri et al. (2021)	US	DJSI US and FTSE4Good VS S&P500 and FTSE USA	Return	Weekly	TERM	Nowcasting
8	Ekinci and Bulut (2021)	Istanbul	Bist100	Return	Weekly	TICKER	Forecasting and nowcasting
9	Smales (2021)	Multiple	G7 G20	Return, volatility	Daily	TERM	Nowcasting
10	Salisu and Vo (2021)	Vietnam	Vietnam Stock Exchange	Return	Weekly	CN	Forecasting
11	Petropoulos et al. (2021)	Multiple	Dax, S&P500, Nikkei 225	Return	Monthly	TERM	Forecasting
12	Fan et al. (2021)	Taiwan	TAIEX	Return	Weekly	CN, INDEX, TERM	Forecasting
13	Sifat and Thaker (2020)	ASIA	ASEAN-5	Return, volatility	Weekly	TICKER AND CN	Forecasting
14	Ding et al. (2020)	US	NASDAQ	Return	Daily	TERM	Nowcasting
15	Ahundjanov et al. (2020)	Multiple	S&P500; DJ; NASDAQ; FTSE100; DAX; CAC40; NIKKEI 225; SZSE; BSE SENSEX	Return	Daily	TERM	Nowcasting
16	Piñeiro et al. (2020)	US	S&P global water index, S&P agriculture livestock index and S&P Energy index.	Return	Monthly	TERM	Nowcasting
17	Nguyen et al. (2020)	Vietnam	HOSE	Return, TV	Weekly	TICKER	Nowcasting
18	Lyócsa et al. (2020)	Multiple	S&P500; FTSE100; NIKKEI225 CACA40; NIFTY50; S&P/TSX DAX; SMI KOSPI and All Ordinaries	Return	Daily	TERM	Forecasting
19	Yoshinaga and Rocco (2020)	Brazil	Ibovespa	Return	Weekly	TICKER	Forecasting
20	Ramos et al. (2020)	Europe	EUROSTOXX50	R, Vol, TV	Weekly	CN	Forecasting
21	Swamy et al. (2019)	India	S&P BSE 500	Return	Weekly	CN	Forecasting
22	Huang et al. (2020)	US	S&P500	Return	Weekly	TERM	Forecasting
23	Kim et al. (2019)	Norway	Norwegian	R, Vol, TV	Weekly	CN	Forecasting
24	Swamy and Dharani (2019)	India	NIFTY 50	Return	Weekly	CN	Forecasting
25	Tan and Taş (2019)	Turkey	BIST	Return	Weekly	TICKER	Forecasting
26	Chen and Lo (2019)	Taiwan	Taiwan top 50 firms	R, Vol, TV	Monthly	CN	Forecasting
27	Desagre and D'Hondt (2019)	Multiple	BEL20; SBF120; AEX25; NASDAQ100; S&P500	Trading Volume	Monthly	TICKER	Nowcasting
28	Padungsaksawas di et al. (2019)	ASIA	Shangai Stock Exchange; BSE; KLSE; Set 50(Thai); ASX20;HSI; Nikkei225; KOSPI; NZX; Straits_time (Singapore)	R, Vol, TV	Weekly	INDEX	Nowcasting
29	Basistha et al. (2019)	US	DJIA; S&P500	Volatility	Weekly	INDEX	Forecasting

30	Nguyen et al. (2019)	Multiple	Emerging markets: Indonesia, Malaysia, Philippines, Thailand, and Vietnam listed stocks	Return	Yearly	TICKER	Nowcasting
31	Škrinjarić (2019)	CESEE region	Hungary, Croatia, Ukraine, Slovenia, Poland, Bosnia Czech Republic, Slovakia, Bulgaria and Serbia stock index	Return, volatility	Monthly	INDEX	Forecasting
32	Oliveira-Brochado (2019)	Portugal	PSI-20 TR	R, TV	Weekly	TERM	Forecasting
33	Chronopoulos et al. (2018)	US	S&P500	Return, volatility	Daily	INDEX	Forecasting
34	Hu et al., (2018)	US	S&P500; DJIA	Return	Daily	INDEX	Forecasting
35	El Ouadguiri and Peillex (2018)	US	DJ Islamic Market US index, FTSE Shariah USA Index, DJ US total stock market index and FTSE USA index.	Return	Weekly	TERM	Nowcasting
36	Pereira et al. (2018)	Multiple	TSX composite index; IPC Mexico; DJ; Nikkei 225; PSE index; DAX; FTSE; CAC40; FTSE min index; ISE; PSI 20; ATHEX 20; Ibovespa; Austral Securities Exchange.	Return, volatility	Daily	TERM	Nowcasting
37	Perlin et al. (2017)	Multiple	S&P500; FTSE100; S&P/ASX200 S&P/TSX	R, Vol, TV	Weekly	TERM	Forecasting
38	Lobão et al. (2017)	Europe	EUROSTOXX50	Return	Monthly	CN	Nowcasting
39	Adachi et al. (2017)	Japan	Mothers and JASDAQ	R, TV	Weekly	CN	Nowcasting
40	Ramos et al. (2017)	Brazil	Ibovespa	R, Vol, TV	Weekly	TICKER AND INDEX	Forecasting
41	Chen (2017)	Multiple	67 Countries	Return	Monthly	TICKER AND CN	Forecasting
42	Moussa et al. (2017)	France	CAC40	Trading Volume	Weekly	CN	Forecasting
43	Tang and Zhu (2017)	US	US ADR	Return	Daily	TICKER	Forecasting
44	Moussa et al. (2017)	France	CAC40	Return, volatility	Weekly	CN	Nowcasting
45	Bijl et al. (2016)	US	S&P500	Return	Weekly	CN	Forecasting
46	Dimpfl and Jank (2016)	US	DJIA	Volatility	Weekly	INDEX	Forecasting
47	Tantaopas et al. (2016)	Asia-Pacific Region	Equity market indices in the Asia-Pacific Region. Developed countries: Australia, Hong Kong, Japan, New Zealand, Singapore, and S. Korea; Developing: China, India, Malaysia, and Thailand.	R, Vol, TV	Weekly	INDEX	Forecasting
48	Arditi et al. (2015)	US	DJIA, NASDAQ100	Return	Daily	TICKER	Forecasting
49	Heiberger, 2015	US	S&P100	Return	Weekly	CN	Nowcasting
50	Hamid and Heiden (2015)	US	DJIA	Volatility	Weekly	INDEX	Forecasting
51	Kristoufek (2015)	US	DJIA	R, Vol, TV	Daily	CN	Nowcasting

52	Takeda and Wakao (2014)	Japan	Nikkie225	R, TV	Weekly	CN	Nowcasting
53	Vozlyublennaiia (2014)	US	DJIA; S&P500; NASDAQ100	Return, volatility	Weekly	INDEX	Forecasting
54	Preis et al. (2013)	US	DJIA	Return	Weekly	TERM	Forecasting
55	Aouadi et al. (2013)	France	CAC40	Vol, TV	Weekly	CN AND INDEX	Nowcasting
56	Kristoufek (2013)	US	DJIA	Return	Weekly	TICKER	Nowcasting
57	Dzielinski (2012)	US	S&P500	Return, volatility	Weekly	TERM	Nowcasting
58	Vlastakis and Markellos (2012)	US	S&P500	Vol, TV	Weekly	CN	Nowcasting
59	Da et al. (2011)	US	Russell 3000	Return	Weekly	TICKER	Forecasting
60	Bank et al. (2011)	Germany	Xetra-listed stocks	R, TV	Weekly	CN	Forecasting
61	Joseph et al. (2011)	US	S&P500	R, TV	Weekly	TICKER	Forecasting
62	Preis et al. (2010)	US	S&P500	TV	Weekly	CN	Forecasting

Source: Based on Ayala et al. (2024)

Table 2. Literature review classification based on financial variable, Keyword, Region and Frequency.

Factor	n° articles	Category	Share
Financial Variable	54	R	87,10%
	24	Vol	38,71%
	20	TV	32,26%
Keywords	21	CN	33,87%
	17	TICKER	27,42%
	16	TERM	25,81%
	14	INDEX	22,58%
Region	24	US	38,71%
	12	Europe	19,36%
	14	Asia	22,58%
	2	South America	3,23%
	10	Multi-country	16,13%
Frequency	11	Daily	17,74%
	41	Weekly	66,12%
	9	Monthly	15,52%
	1	Yearly	1,63%

Source: Based on Ayala et al. (2024)

3.3.1. GSVI Building criteria: Keyword Alternatives and Data Frequency

3.3.1.1. *GSVI Keyword Alternatives*

One of the most important and controversial criteria in the GSVI literature is the keyword used to create the index. In the reviewed articles, four classifications can be identified: Company name (CN), Ticker symbol (TICKER), related term (TERM), and index name (INDEX). According to the papers reviewed, Company name is the most used keyword, followed by ticker, related terms and index.

There are contrasting perspectives on selecting less ambiguous keywords for the GSVI construction. Da et al. (2011) and Joseph et al. (2011) pointed out two major challenges in using company names as keywords in the GSVI. Firstly, the search term does not always directly reflect investment intent, as it could be related to other purposes, such as online shopping or finding store locations. Secondly, company names may appear in various forms, including abbreviations, which complicates the process. To overcome these issues, they propose using ticker symbols, since those searching for ticker codes are more likely to have investment intentions, which would improve the accuracy of the GSVI data. Broder (2002) also noted that while search traffic data provides a general idea of what the user is searching for, it may not capture the precise "need behind the query." However, Kristoufek (2013) found that using stock tickers as a dataset sample poses challenges due to the difficulty in achieving a high frequency for some terms, which can reduce the scope of the dataset. Similarly, Tan and Taş (2019) manually excluded certain ticker symbols, as they had broader meanings that coincided with common Turkish words, thereby introducing noise and limiting the quality of the dataset.

Although there are discrepancies in the use of firm names, a considerable number of articles have constructed the GSVI using company name data. Vlastakis and Markellos (2012) argue that the noise highlighted by Da et al. (2011) and Joseph et al. (2011) is random and does not significantly impact the GSVI data sample. To address the second potential issue, several authors have proposed alternative keyword selection methods. Ramos et al. (2020), Bijl et al. (2016) and Preis et al. (2010) adapted the use of company names by simplifying the search term, often removing irrelevant parts like "Inc." Takeda and Wakao (2014) and Adachi et al. (2017), developed a comprehensive method specifically for Japanese companies,

excluding abbreviations with general meanings. Moussa et al. (2017a) utilized a service called "AdWords" to enhance keyword searches by limiting the terms to the "investment" category. Additionally, *Google Trends* enables region-specific filtering of query data, which reduces bias caused by variations in company names across different languages.

Among the papers reviewed, several used related terms as keywords. Oliveira-Brochado (2019), Perlin et al. (2017) and Preis et al. (2013) employed financial terms to construct the GSVI. These studies not only used financial terms to measure their relationship with stock market performance but also incorporated terms relevant to the subject under study. Chundakkadan & Nedumparambil (2022), Smales (2021), Ahundjanov et al. (2020), Ding et al. (2020) and Lyócsa et al. (2020) focused on search terms related to COVID-19 to capture investor attention to the pandemic and its effects on stock markets. Dzielinski (2012) used the term "economy" to assess economic uncertainty through the GSVI, alongside volatility and returns of the S&P 500. Other studies followed similar term selection methods, such as Piñeiro et al. (2020) for the water sector index, El Ouadghiri and Peillex (2018) for measuring investor attention to Islamic terrorism, Ouadghiri et al. (2021) for the climate change sector, and Pereira et al. (2018) for the global impact of the "Donald Trump" effect. Additionally, some articles used index names as keywords. Aouadi et al. (2013) suggested that trading volume and investor attention correlate more strongly at the market level, as measured by index names, compared to firm-specific data. In line with this, Ghaddab et al. (2024), Sukamulja and Thedora (2023) and Tantaopas et al. (2016) examined index-level data, arguing it offers a more comprehensive representation than individual firms.

3.3.1.2. *GSVI data frequency*

One of the main features to consider about *Google Trends* is the frequency with which it provides search volume data. *Google Trends* provides data at different frequencies based on the length of the analysis period, which offers researchers flexibility in their approach. Since its inception in 2004, *Google Trends* has made data available daily for analysis periods of up to nine months. For longer periods, the frequency of the data shifts to weekly or monthly intervals, depending on the duration of the timeframe being studied. As such, researchers must carefully select the appropriate frequency to align with their study objectives, as this can

significantly impact the findings. However, the research community has not thoroughly investigated the implications of choosing one frequency over another, leaving a gap in the literature. Based on the reviewed sample, the frequency of the search volume can be categorized as weekly, daily, monthly, or yearly.

Ekinci and Bulut (2021) argue that utilizing weekly data may not effectively capture the dynamics driving market movements. Consider a scenario where an increase in the GSVI is observed on a Monday, and stock returns rise on Wednesday. In this case, one might deduce a causal link between the GSVI increase and the subsequent rise in returns. Conversely, if stock returns rise on Thursday and the GSVI increases on Friday, the causal interpretation would be reversed, highlighting the complexities in identifying temporal relationships when weekly data are used. Similarly, Hamid and Heiden (2015) caution against using daily frequency data for the GSVI when forecasting market volatility, arguing that it may not provide sufficient predictive power or context. Despite these concerns, approximately 40% of studies employing daily frequency have incorporated volatility as a critical financial variable in their analyses.

Adjusting the analysis period can allow researchers to refine the frequency of their data. For instance, narrowing the analysis to shorter timeframes makes it possible to examine patterns and trends with greater precision. Nevertheless, the majority of reviewed studies (66.1%) opted for weekly frequencies for both search queries and financial metrics. This choice reflects a balance between capturing market dynamics and managing data complexity. Daily and monthly frequencies were utilized with nearly equal frequency, while yearly data was employed in only one study (Nguyen et al. 2019). In that case, the author initially gathered monthly data and later aggregated it to create annual figures for their analysis.

Notably, studies focusing on daily data collection have often examined recent, event-driven phenomena, such as the COVID-19 pandemic. Research by Chundakkadan and Nedumparambil (2022), Smales (2021), Ahundjanov et al. (2020), Ding et al. (2020), and Lyócsa et al. (2020) explored how COVID-19-related search trends influenced investor behaviour and stock market reactions. Given the short timeframe of these studies (typically less than a year), daily data collection was deemed appropriate. For such time-sensitive research, higher-frequency data provided actionable information into the rapid shifts in market sentiment and

investor attention, whereas longer frequencies like weekly or monthly were less applicable.

Ekinci and Bulut (2021) further caution that selecting an inappropriate data frequency can result in information leakage or misinterpretation of market signals. Asset prices in financial markets can change rapidly, with significant boom-and-bust cycles unfolding within seconds. Therefore, using high-frequency data, such as daily or even intraday data, allows researchers to capture these swift market movements. Higher-frequency data not only increases the number of observations available but also enhances the robustness of the findings. Araújo et al. (2015) emphasize that acquiring high-frequency data is critical for accurately reflecting the volatility and dynamics of asset markets. This approach enables researchers to better account for fluctuations, reduce noise in their analyses, and improve the overall reliability of their results.

While weekly frequencies dominate the literature, daily data is increasingly relevant in event-driven studies, particularly in contexts requiring high temporal resolution, such as pandemic-related market analysis. However, the broader research community would benefit from a more comprehensive exploration of how frequency choices impact the interpretation and reliability of financial data.

3.3.2. GSVI studies by market region

Another relevant criterion is the geographical focus of the studies. Different stock markets exhibit distinct characteristics, such as varying levels of liquidity and market capitalization, all of which can influence the observed relationship between search activity and market performance. Based on the reviewed papers, the regions of study are US, Europe, Asia, South-America and multi-country analysis.

The majority of the reviewed investigations rely on data from the US stock markets, which are among the most well-structured and provide more extensive data compared to other countries. Despite Europe's strong academic output, only a small portion of the reviewed studies (roughly a fifth) examine the stock markets of European countries. The S&P 500 stands out as the most frequently analysed stock index, whether in isolation or as part of cross-country comparisons with other indices. Together with the DJIA these two indices account for about half of the

research conducted. Given that the US stock market is one of the most actively traded globally, companies listed there provide valuable understanding of how market indicators, such as returns, trading volume, and volatility, respond to various types of news.

During periods of market activity, investors tend to demonstrate increased interest and motivation to engage with stock markets, leading to a surge in information-seeking behaviour (Yoshinaga and Rocco 2020). As a result, researchers are able to gather stronger and more reliable samples from these markets, which effectively reflect investor attention, interests, and behaviour.

In addition, eleven cross-country studies offer stronger evidence of the GSVI's effectiveness in measuring investor attention and forecasting across a broader range of markets. However, one limitation in these studies is the uneven distribution of Internet access in the regions covered by the search traffic data (Sifat and Taker 2020). Another regional constraint is the presence of alternative search engines, such as Baidu, which is more commonly used in Asia than Google. As a result, studies conducted in Asian markets face limitations in the volume of search queries available. Despite these challenges, Asian markets still account for about one-fifth of the studies reviewed.

Similarly, European financial markets, led by the EUROSTOXX50, along with other major indices like the CAC40 and FTSE100, contribute to another fifth of the research examined. Lastly, the South American region is represented by two studies focused on the Brazilian stock markets (Yoshinaga and Rocco 2020; Ramos et al. 2017), one of the BRICS economies and a key player among emerging markets globally.

3.3.3. GSVI and financial variables: Return, Volatility and Trading Volume.

The previous sections reviewed the literature related to methodological criteria, such as data frequency, keyword construction alternatives for the GSVI, and the variability of results across analysed markets. This thesis delves deeply into the capacity of Google searches to explain current and future stock market movements. Furthermore, the research examines individually the impact of the

GSVI on stock returns, volatility, and trading volume, thereby achieving a comprehensive understanding of stock market behaviour. For this reason, the literature review in this section has been structured in the same manner.

3.3.3.1. *GSVI and Return*

Stock return is key variable in explaining the potential behavioural signals of stock markets. Most of the studies reviewed that analyse the effects of investor attention on stock market behaviour (87%) use it as their main variable. However, literature is inconclusive about the relation between the GSVI and stocks Returns.

Da et al. (2011) studied a sample of Russell 3000 stocks from 2004 to 2008, using the GSVI to measure investor attention. Their findings suggest that an increase in search volumes predicts higher stock prices over the next two weeks, followed by an eventual price reversal within the year. Similarly, Joseph et al. (2011) developed a trading strategy based on ticker search volumes and concluded that search intensity reliably predicts abnormal stock returns. Bank et al. (2011) confirmed these results, demonstrating that portfolios based on Google search volume (using company names) could be shortened, with similar outcomes observed in the German stock market. Nguyen et al. (2019) and Salisu and Vo (2021) provided evidence of a sustained, positive relationship between search volume and stock returns in the Vietnamese stock market. This finding aligns with the research by Adachi et al. (2017), who examined 208 firms from Japanese start-up market, revealing a positive relationship between Google search volumes and stock returns.

The positive relationship between the GSVI and current and future stock returns found in the aforementioned studies is consistent with the propositions stemming from the work of Barber and Odean (2008), as the results improve short-term buying pressure accompanied by an increase in stock prices. In accordance to Adachi et al. (2017), their results confirm the PPH (Barber and Oden, 2008), remarking that an increase in search volume is associated with subsequent positive stock returns. However, the authors did not find sufficient evidence to support the price reversal, thus arguing that the relationship persisted in the long term. The authors suggested that their results might be influenced by the type of companies in the sample, as smaller firms tend to have higher returns due to the impact of new information. A similar long-term relationship was found by Nguyen et al. (2019), although the intensity of the effect was greater for larger companies than for

smaller ones. However, the author noted that these results could be attributed to the immaturity of the Vietnamese market.

Other studies, including those by Aouadi et al. (2013) in France, Swamy et al. (2019) in India, Takeda and Wakao (2014) in Japan, Tan and Taş (2019) in Turkey and Lobão et al. (2017) in Europe, have also found positive relations between Google search volumes and stock returns, supporting the idea that investor attention, as reflected in search volume, can predict stock market movements across different regions.

Bijl et al. (2016) conducted an analysis using data from 2008 to 2013, concluding that Google search volume, as an indicator of investor attention, is negatively related to stock returns. Although the coefficients were statistically significant, the authors noted that the impact of this relationship was relatively small. Based on their findings, they developed a trading strategy that, when accounting for transaction costs, proved to be unprofitable. This could suggest that while Google search volume may provide some insight into market movements, its practical application in trading strategies is limited. In contrast, Perlin et al. (2017) reinforced the idea that Google searches have predictive power over stock markets. They focused on financial terms such as "*stock*" and "*finance*" as keywords and found that an increase in searches for the term "*stock*" in the weeks prior was negatively correlated with stock returns. Moreover, the authors designed a portfolio model and discovered that the predictive ability of Google searches was more pronounced during financial crises, highlighting the potential for search volume data to be particularly valuable during periods of market stress. These findings suggest that while the overall relationship between search volume and returns may vary, the context, such as market conditions, can significantly influence the strength of the predictive power of Google searches.

Kim et al. (2019) conducted a study on the Norwegian market but observed no significant relationship between search intensity and stock returns, either in the present or the future. The authors attribute these results to the small sample size used in their analysis, suggesting that it may have limited the ability to detect meaningful patterns. Similarly, Ramos et al., (2019), in their examination of the European market between 2004 and 2011, did not identify a significant connection between Google searches and stock returns. Cheraghali et al. (2023) reached comparable conclusions, noting the absence of a nowcasting relationship between

Excess returns and the GSVI. However, they did recognize its predictive power, observing that increased investor attention in the prior period was negatively correlated with stock returns.

Other studies that have identified a negative relationship between investor attention and financial returns include Piñeiro et al. (2020) in the US, Yoshinaga and Rocco (2020) in Brazil and Heiberger et al. (2015) in Germany. Additionally, Akarsu and Söer (2022) conducted a cross-country analysis encompassing 31 countries across the Americas, Asia-Pacific, and Europe, concluding that the impact of investor attention on stock returns varies significantly between countries.

3.3.3.2. *GSVI and Volatility*

Another widely used financial metric in the literature to assess market behaviour is volatility. It captures the magnitude and frequency of price fluctuations, serving as a key indicator of market risk and uncertainty. Vlastakis and Markellos (2012) show that investor attention, measured by the GSVI, significantly affects volatility at both the firm-specific level, tracked through company names, and the market-wide level. Dimpf and Jank (2016) found that high volatility precedes increased investors' information demand. Using daily market-level data with the keyword "DOW," they demonstrate that Google searches are strongly contemporaneously correlated with market volatility. Search queries serve as an indicator of future volatility, with increased investor attention to the stock market leading to higher volatility levels. Additionally, incorporating the GSVI into forecasting models significantly improves their accuracy, especially during times of elevated market volatility.

The effect on volatility has also been examined in the context of socioeconomic events. For instance, Pereira et al. (2018) found that investor behaviour is influenced by events related to Donald Trump's presidency. Using the search term "Donald Trump," the authors investigated its effect over a two-year timeframe (on a daily frequency) starting from the day Trump announced his candidacy. Their findings reveal a positive correlation between search activity and market volatility in countries such as Mexico, Japan, and Australia.

In Asia, Chen and Lo (2019) studied the relationship between online search activity and stock market performance for Taiwan's top 50 firms, identifying a positive association between the GSVI and stock volatility. Similarly, Ramos et al.

(2020) analysed European markets by investigating the effect of investor attention on the EUROSTOXX50 index. They concluded that an increase in search queries is followed by a short-term rise in volatility, which reverses after two weeks.

Recent studies, such as Smales (2021) and Chundakkadan & Nedumparambil (2022), explore the impact of COVID-19 news on investor sentiment and its influence on stock market behaviour. Using a multi-country sample, they find that pandemic-related sentiment has led to increased market volatility, with higher internet search activity related to COVID-19 being associated with greater stock volatility. Similarly, González-Velasco and González-Fernández (2023) build their research on the premise that investor decisions are shaped by sentiment, including fear. Their study examines the effect of the fear response to the COVID-19 pandemic on stock market volatility, highlighting its significant role in driving market fluctuations.

In conclusion, the literature consistently supports the existence of a positive relationship between search volume and volatility. This suggests that when investors are more focused on specific stocks or the market as a whole, indicated by increased online search activity, volatility tends to rise.

3.3.3.3. *GSVI and Trading Volume*

Finally, the effect of investor attention, as measured by information demand, is related to the trading volume.

Bank et al. (2011), in their study of the German market using Xetra-listed stocks, examined the relationship between internet search activity and trading activity through stock liquidity. The authors aimed to assess whether this relationship affected the level of asymmetric information in the market. The study concluded that an increase in the GSVI is positively correlated with a rise in trading activity. When investors engage in more online searches, they become better informed, thereby reducing asymmetric information. Similarly, Joseph et al. (2011), in their analysis of companies listed on the S&P 500 index, found that over a one-week period, higher online search intensity can predict increased trading volumes. Moreover, this relationship is even stronger for stocks that are more difficult to arbitrage, suggesting that investor attention, as indicated by search activity, has a more pronounced impact on stocks with higher market inefficiencies.

Following the same approach as Joseph et al. (2011), Takeda and Wakao (2014) tested the predictive power of the GSVI on the Asian Nikkei 225 index. They claim that the higher the intensity of the GSVI, the greater the abnormal trading volume reported. In a similar vein, Nguyen et al. (2020) studied the HOSE index by conducting a portfolio analysis based on the GSVI intensity, finding that an increase in search attention positively affects abnormal trading volume.

Aouadi et al. (2013) also support the positive relationship, although they found a stronger correlation between the GSVI and trading volume that was more market-related than firm-specific. Building on these results, Tantaopas et al. (2016) conducted their research at a broader index level, arguing that it provides a better reflection of individual investors' investments, rather than focusing on specific-firm stocks.

The findings of Chen and Lo (2019) and Kim et al. (2019) align closely, as both studies conclude that the GSVI is significantly positively correlated with abnormal trading volume. Furthermore, they provide evidence that GSVI serves as a direct proxy for investor attention.

Building on the previous studies, Desagre and D'Hondt (2019) explored the relationship between investor attention and trading activity in a multi-country analysis of 455 stocks. They concluded that this relationship was positive, but not stronger for purchases than for sales, thus not supporting the PPH. According to this hypothesis, attention influences buying behaviour more than selling behaviour, as investors typically have more options when purchasing than when selling. As a result, an increase in attention (reflected by a rise in the GSVI) is expected to temporarily pressure prices, leading to more traded financial assets, as suggested by Barber and Odean (2008).

Recently, Sukamulja and Thedora (2023) found a significant link between investor attention and trading activity across four out of the five indices they studied. The only exception was the Indonesian market, where no such relationship was observed. According to the authors, these characteristics may limit the impact of attention on trading activity, highlighting the importance of considering market-specific factors when analysing the relationship between attention and trading behaviour.

IV - DEVELOPMENT OF RESEARCH HYPOTHESIS

IV -DEVELOPMENT OF RESEARCH HYPOTHESIS

Investor attention plays a crucial role in financial decision-making, as it influences trading activity, price volatility and market efficiency. This thesis focuses on investigating whether cognitive biases of investors and their limitations in thinking and information influence the European market and quantifying this impact, thereby contributing to the ongoing debate about the efficiency of financial markets.

Previous studies have explored the relationship between search volumes and trading variables, with some focusing on contemporaneous effects (Ghaddab et al., 2024; Ouadghiri et al., 2021; Ding et al., 2020) and others on predictive capabilities (Sukamulja and Thedora, 2023; Salisu and Vo, 2021; Ramos et al., 2020). Given the dynamic nature of financial markets, understanding both current and future interactions between the GSVI and financial variables is crucial.

At this stage, and in order to present the research hypotheses of this thesis with greater precision, it is essential to clarify that the empirical models will incorporate lagged variables for both the GSVI and market-related variables. The explanatory power of the GSVI will be assessed concurrently with the exogenous variables (return, volatility and trading volume), allowing for an examination of how the dependent variables respond to an increase in search intensity (Cheraghali et al., 2023; Piñeiro et al., 2020; Lyócsa et al., 2020; Kim et al., 2019). The definition of the empirical model will be explained deeper in the methodology section.

In line with previous literature (Yoshinaga and Rocco, 2020; Kim et al., 2019; Swamy et al., 2019; Perlin et al., 2017; Bijl et al., 2016), the predictive capacity of Google search traffic is analysed by introducing lagged values of the GSVI variable by one and three weeks. This methodological approach facilitates the assessment of whether the GSVI exhibits predictive power in forecasting future values of financial variables.

Moreover, the empirical study and hypotheses are structured around the key financial variables. Accordingly, each research hypothesis will include two sub-hypotheses, depending on whether the aim is to examine the contemporaneous relationship or the predictive capacity of Google searches. Based on the previous premise, this thesis formulates the following research hypotheses:

GSVI and stock returns

The first two hypotheses are developed to examine the relationship between the GSVI and stock returns. Research literature supports the predictive ability of GSVI on stock returns (Huang et al., 2020; Lyócsa et al., 2020; Oliveira-Brochado, 2019; Swamy et al., 2019; Chronopoulos et al., 2018). However, findings regarding the direction of the relationship between increased investor attention and stock returns remain inconclusive.

Yoshinaga and Rocco, (2020), Huang et al. (2020) and Bijl et al. (2016), argue that higher search traffic is negatively associated with stock returns, as increased attention and popularity can drive stock prices away from their fundamental value. Nevertheless, the overall impact appears to be small. Alternatively, authors such as Ouadghiri et al. (2021) and Da et al. (2011) in the US; Sifat and Thaker (2020), Nguyen et al. (2020) and Adachi et al. (2017) in Asian region; Swamy et al. (2019) in Indian markets or Ramos et al. (2017) in Brazil argue that an increase in search activity is positively associated with stock returns.

These studies support the PPH (Barber and Odean, 2008), which suggests that a sudden and substantial increase in demand for a security result in a temporary rise in its price. According to this theory, investors adopt a buying behaviour when actively seeking information, leading to a surge in buying demand, which exerts upward pressure on prices, causing them to rise. This demand of information can be captured by Google searches (Chen et al., 2022; Adachi et al., 2017), which offers a quantifiable way to anticipate potential price pressure events in the market. On account of that, this thesis expects that European markets perform on a similar way, presenting the hypotheses as follow:

RH_{1a}: GSVI is positively related with stock returns in European markets.

RH_{1b}: GSVI predicts stock returns in European markets.

GSVI and volatility

The relationship between search intensity (GSVI) and volatility is also tested. Vlastakis and Markellos (2012) studied the US market, concluding that the GSVI significantly impact volatility in both firm-specific and market level. Similar results were obtained by Chundakkadan & Nedumparambil (2022); Pereira et al. (2018) and Dimpf and Jank (2016), who observed the positive relation between searches and volatility measure. Their findings suggest that heightened search activity, driven by increased investor attention, correlates with high volatility. As investors search for information, they are typically responding to perceived uncertainty or the anticipation of potential market events.

Additionally, the GSVI might be also considered as a potential predictor of future price fluctuations. This increased demand for information reflects a higher level of uncertainty, leading to more significant price fluctuations. Basistha et al. (2019) concluded that search volumes have the ability to predict volatility in commodity markets. Ramos et al. (2020) found that an increase on attention is followed by a small increase in volatility in the components of EUROSTOXX50 index. Accordingly, the second group of hypotheses is presented as follow:

RH_{2a}: *GSVI is positively related with stock volatility in European markets.*

RH_{2b}: *GSVI predicts volatility in European markets.*

GSVI and trading volume

Lastly, investors' attention is related with trading activity. Investors are net buyers of attention-grabbing stocks and an increase in the attention demand paid to the market will positively pressure the stocks liquidity, hence, assets trading volume (Nguyen et al., 2020; Adachi et al., 2017). The GSVI, as a measure of investor interest, reflects this dynamic. When the demand for information increases, driven by factors such as market news or emerging trends, it typically signals greater

investor involvement, which leads to higher trading volumes as investors act on the information they gather (Desagre and D'Hondt, 2019; Perlin et al., 2017)

Several studies have confirmed this relationship. Ramos et al. (2020); Takeda and Wakao (2014) and Preis et al. (2010) found that the GSVI has a positive impact on trading volume. According to Kim et al. (2019), an increase in search intensity suggests higher investor interest, rising market participation. Furthermore, the authors found that information demand can predict future trading volumes, even after adjusting for control variables.

Based on this, the last two hypotheses are formulated as follow:

RH_{3a}: *GSVI is positively related with stock trading volume in European markets.*

RH_{3b}: *GSVI predicts trading volume in European markets.*

V- METHODOLOGY

V - METHODOLOGY

5.1. DATA COLLECTION

The study selects companies that have been included in the EUROSTOXX50 index during the years of the sample period. Data spans the period from June 30th, 2014 to June 30th, 2024. Data from 2013 and part of 2012 was also necessary to build the financial variables and standardize the some of the variables. The raw search queries volume was retrieved from *Google Trends* and financial data such us daily open, close, high, low and adjusted prices values and trading volume from EIKON platform⁴. Market dynamics are evaluated through three financial variables: Return, volatility and trading volume. Investor attention is proxied by the GSVI, estimating it with three different methodologies. In the following sections 4.1.1 and 4.1.2 the variables are presented.

5.1.1. Constituents Data

To achieve a representative sample of the European market, companies that have been part of the EUROSTOXX50 index over the past decade are used as references. The stock index EUROSTOXX50 consists of 50 large-cap and liquid companies that are leaders in the respective industries. It is designed to provide a blue-chip representation of super sector leaders in the Eurozone. The index includes firms from Germany, Netherlands, France, Belgium, Spain, Italy, UK, Finland, Ireland, Switzerland, Denmark and Sweden, being traded in different stock exchanges such as Bolsas y Mercados Españoles, Euronext, Deutsche Borse or London Stock Exchange.

In order to construct the data sample of companies, the following pre-selection criteria are defined:

1. Following Ramos et al. (2019), it is included companies that have joined the EUROSTOXX50 index at least once during the period from 2014 to 2024. In this manner, a more comprehensive perspective can be attained,

⁴ <https://eikon.refinitiv.com/>

viewing it as a global market rather than merely an index. It is considered that if a company has been included in the index at any point, it indicates that it is representative of the European market. Therefore, including such a company enhances the robustness of the data sample.

2. Companies with more than two years of trading data. It ensures that the companies in the sample have a minimum track record of market activity, allowing for a more robust analysis of their financial behaviour
3. Companies that have remained in the index for more than one month.

The sample of companies obtained after applying the inclusion criteria consists of 80 firms. However, a significant challenge arose due to the inherent limitations of Google Trends data, as search volume information is not uniformly available for all companies. For some firms, search volume data is unavailable, as Google Trends reports a value of zero when the number of searches is not representative. The initial number of search volume observations for the 80 companies is 28,695. However, due to methodological criteria, 28 companies were dropped, resulting in a final sample of 52 firms. After excluding these 28 companies, the number of observations is reduced to 27,481, meaning that the excluded observations represent only a small proportion of the initial dataset.

This sample size ($n=52$) is consistent with similar studies in the literature. For instance, Ramos et al. (2019) conducted a study on European markets with a final sample of 36 firms. Likewise, Kim et al. (2019) analysed companies listed on the Norwegian OBX stock index, reducing their sample from 40 initial firms to a final set of 28.

Appendix 1 shows the list of companies included in the data sample. Table 3 illustrates the distribution of companies by country, with France, Germany, and the UK emerging as the nations with the highest business representation in the European market. Table 4 provides an overview of the sectoral classification of the constituents, highlighting that the financial, consumer cyclical, and technological sectors constitute the largest share within the selected sample.

Table 3. Country distribution of EUROSTOXX50 constituents.

COUNTRY	N	SHARE
France	14	26,9%
UK	13	25%
Netherland	8	15,4%
Spain	6	11,5%
Italy	5	9,6%
Germany	2	3,8%
Switzerland	1	1,9%
Finland	1	1,9%
Belgium	1	1,9%
Ireland	1	1,9%
Total	52	100,0%

Source: Own elaboration based on <https://eikon.refinitiv.com/>

Table 4. Sector distribution of EUROSTOXX50 constituents.

SECTOR	N	SHARE
Financials	12	23,1%
Consumer cyclicals	7	13,5%
Technology	10	19,2%
Industrials	4	7,7%
Consumer non-cyclicals	7	13,5%
Healthcare	3	5,8%
Basic materials	3	5,8%
Utilities	4	7,7%
Energy	2	3,8%
Total	52	100,0%

Source: Own elaboration based on <https://eikon.refinitiv.com>

5.1.2. *Google Trends Data*

5.1.2.1. *Keyword selection*

The choice of keyword to construct the GSVI is crucial for establishing the model. According to Ayala et al. (2024), four main keywords have been used to determine Google search frequencies: company names, stock ticker, index name, and economic terms.

Company name is the most commonly used keyword for constructing the GSVI (Ayala et al., 2024). Using company names as a primary keyword is effective because they are widely recognized, frequently searched by investors and consumers, and serve as a direct identifier of a firm's market presence. However, most of the authors proposed various modifications of company names, such as adaptations, simplifications, or the removal of irrelevant components like "Inc." (Ramos et al., 2020; Bijl et al., 2016; Preis et al., 2010).

Despite being the most used keyword category, Da et al. (2011) and Joseph et al. (2011) identified two problems with this keyword category. First, the multiple meanings of a company name (e.g., "Orange" can be a color, a fruit or a company). Second, it is difficult to identify the need behind the search. For instance, if we look for the word "Santander," the Spanish banking company, we cannot determine if the person needs a physical bank, is looking for the location of the Spanish city "Santander," or wants some financial information about the bank.

Nguyen et al. (2020) and Desagre and D'Hondt (2019) use the ticker symbol as a keyword, arguing that people may search for it primarily for investment reasons. The ticker symbol is the identification keyword for a specific company in the stock market. Hence, there would not be any other reason for searching it rather than investment purpose.

The main issue associated with the utilization of ticker symbols is that it can be tough to find data for some firms. This suggests two potential interpretations: firstly, that the acquired data may lack representativeness, leading to its exclusion from the sample data (Ramos et al., 2020); alternatively, a zero-value result in searches may imply less investor attention toward that company (Adachi et al., 2017).

Recently, deHaan et al. (2023) studied the measurement errors when using stock tickers as Google search queries. They concluded that using the combination of "ticker-stock" (e.g., "SAN stock") as search terms performed better in tests than using just the ticker symbol. As a first attempt, deHaan et al. (2023) was followed to build the GSVI data sample. However, the majority of the searches returned a value of 0.

Finally, following Da et al. (2011) it is used stock ticker as keyword for search queries. The stock tickers are adjusted removing the market denomination. *Google Trends* provides the option to filter by activity channel (Website, YouTube, Google Shopping, news, etc.). Following Kim et al. (2019), it is downloaded the ticker symbol searched in website category.

Lastly, Google Trends allows users to select the geographical scope of the queries. This study opted to download searches conducted worldwide for several reasons. Firstly, the companies analysed are globally recognized, with substantial market capitalization and size, which means their stocks are traded internationally rather than being confined to the country where their headquarters are located. Secondly, financial markets today operate in an increasingly globalized environment, where investor interest and trading activities are not limited by national or regional boundaries. Given this reality, restricting searches to a single European country or region would not accurately reflect the broader market dynamics in which these firms operate. A global perspective ensures a more comprehensive and representative analysis of investor interest and market trends.

5.1.2.2. **GSVI**

Investor attention is not a direct measurable variable as it is price or trading volume. As information flow is not directly observable, it needs a proxy variable to quantify it (Vlastakis and Markellos, 2012). The research community has approximated different indirect ways for measuring investor interest to particular stocks or markets such as unusual trading volume (Wright and Swidler, 2023; Li et al., 2011), social media (Clarke et al., 2020; Li et al., 2016;), or news (Hou et al., 2009). Da et al., (2011) proposed a novel proxy variable denoted as the GSVI for measuring investor attention, which reflects the frequency of search queries of a specific keyword via *Google Trends*. Accordingly, in this thesis, the GSVI will be used as a measure of investor attention based on search activity.

Google Trends uses a normalized approach to represent the search volume data on a range from 0 to 100, being 100 the highest volume of search or maximum interest during the specified period. If the value is 0, there are not enough searches about the keyword and hence it is interpreted as no interest or popularity.

When looking for a term in *Google Trends*, it will provide the search volume based on different uses of the word (Perlin et al., 2017). For instance, the search frequency of the word *milk* will include all the searches related to other variant as *skimmed milk*, *oat milk* or *lactose-free milk*. Thus, one of the predicted constraints is to know the final purpose of the search.

Google Trends also helps to identify seasonal trends and cyclical patterns in search behaviour. Figure 5 illustrates an example of how the tool presents time series data. When looking for *Christmas* in *Google Trends*, it can be showed how the search volume for *Christmas* spikes in December each year.

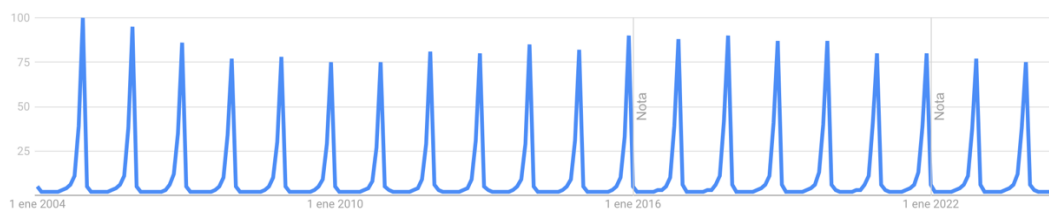


Figure 5. Search intensity of the term *Christmas* from 2004-2024.

Source: <https://trends.google.com/trends/>

Another specificity of the tool is that it allows filtering by geographic area (worldwide or single country), by category⁵ (Arts & Entertainment, Auto, Beauty & Fitness, Books & Literature, Business & Industrial, Computers & Electronics, Finances, etc.) and subcategories, and by activity channel (Website, YouTube, Google Shopping, news, and Google Image) making the data more relevant and actionable for their particular needs.

The frequency for search volumes depends on the length of the period selected. For periods up to 270 days, daily data is available. If the chosen period

⁵ The full list of categories and subcategories is available at <https://trends.google.com/trends/>

exceeds five years, the tool offers the data on a monthly basis. On the other hand, for shorter periods, the frequency can be obtained on a weekly basis. Hence, the frequency of the data provided by the search engine decreases as the interval length increases. This study uses weekly volume data, being consistent with others research methodologies (Akarsu and Süer 2022 ; Ouadghiri et al. 2021).

The raw data available to download from *Google Trends* cannot be directly used and some transformations are required. ~~must be transformed~~. According to Bustos and Pomares-Quimbaya (2020), this raw data is classified as unstructured inputs as it is necessary some text-mining techniques to obtain output models. Furthermore, there are additional considerations to account for when calculating the GSVI.

Firstly, the frequency of the retrieved data. One of the limitations of Google Trends when downloading data is that it only allows access to weekly data for periods of up to five years. For longer time frames, the tool provides monthly values. The time span covered in this study extends from June 30, 2014, to June 30, 2024, and the analysis is based on weekly search data. The selected period is 10 years, but data from at least one more year is needed for the normalization of the variables.

Figure 6 represents the time structure for the GSVI data. The periods have been structured with a one-year overlap between them. For example, period (1) spans from 01/01/2013⁶ to 12/31/2017, and period (2) starts on 01/01/2017, one year before period (1) ends, creating a one-year overlap.

⁶ In the figure the starting date of the period is displayed as the end of the previous year to represent better the year gap.

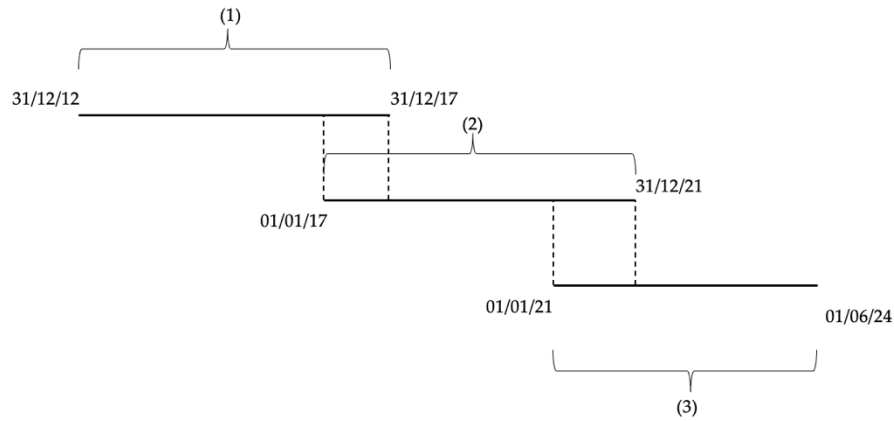


Figure 6. Structure of time-retrieval of raw Google Trends data.

Source: Own Elaboration.

Secondly, the raw data cannot be used directly in the model because the search value depends on the period that it is downloaded. For instance, the search volume of i stock the first week of December 2017 will vary if we take it from period (1) or period (2). To address this time retrieval bias, Da et al. (2011) proposed the Abnormal Google Search Volume Index (AGSVI) to standardize search queries data. This approach eliminates scale bias, enabling a fair comparison of the search terms and remove temporal trends, smoothing the variations in changes. Furthermore, it ensures consistency within the study models and facilitates the interpretation of changes in search volumes in a relative context, enabling comparisons across stocks in cross-sectional analyses. Figure 7 presents the raw search volume data of three companies before the standardization. Figure 8 shows the abnormal search volume of the same three companies.

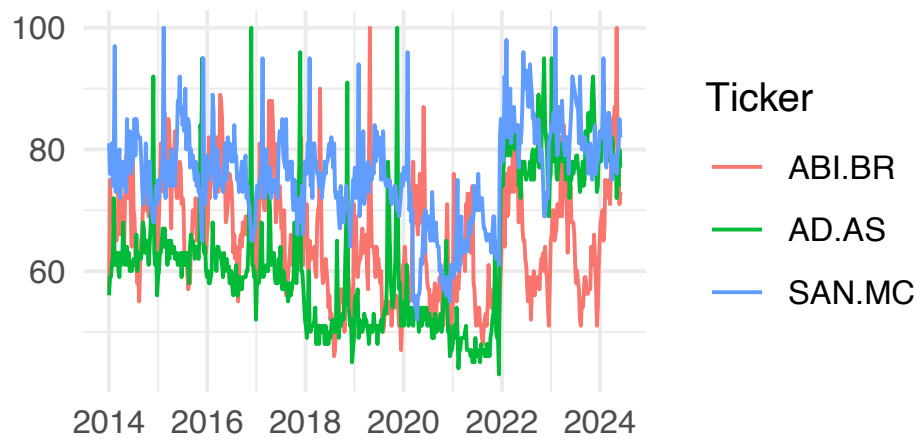


Figure 7. Search volume Raw Data of three companies.

Source: Own Elaboration

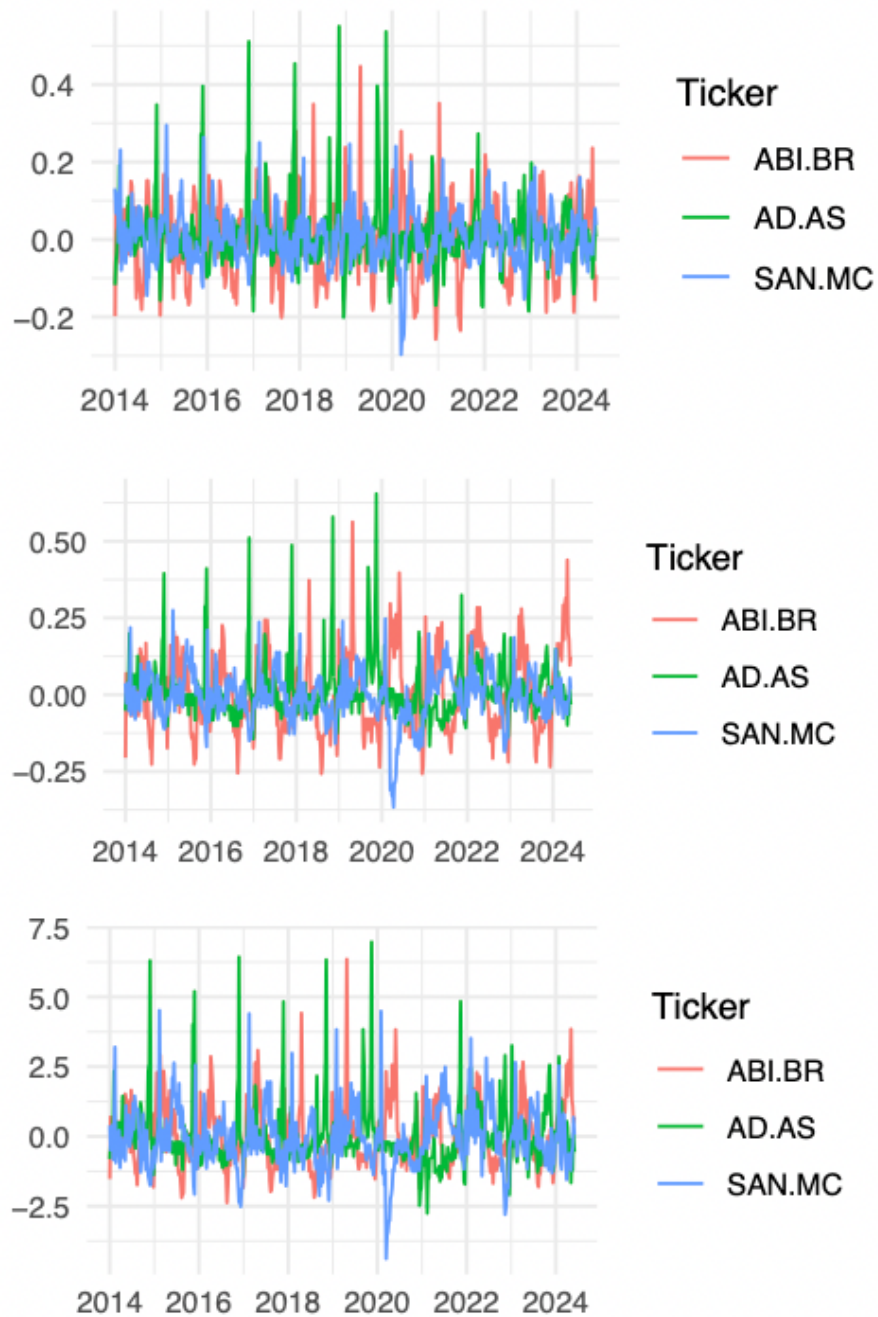


Figure 8. Standardized Google Search volumes using $AGSVI_{DA8,t}$; $AGSVI_{DA52,t}$ and $AGSVI_{BIJL52,t}$

Source: Own Elaboration.

For calculating the AGSVI three methods are followed. Firstly, as in Da et al. (2011):

$$AGSVI_{DA8,t} = \log(GSVI_t) - \log [Med(GSVI_{t-1}, \dots, GSVI_{t-8})] \quad (3)$$

Secondly and as a robustness check, two different time horizons are used as in Bijl et al. (2016), considering the average of the previous 52 weeks and the past year standard deviation:

$$AGSVI_{DA52,t} = \log(GSVI_t) - \log [Med(GSVI_{t-1}, \dots, GSVI_{t-52})] \quad (4)$$

Thirdly, following Bijl et al. (2016):

$$AGSVI_{BIJL52,t} = \frac{GSVI_t - \frac{1}{52} \sum_{i=1}^{52} GSVI_{t-i}}{\sigma_{GSVI}} \quad (5)$$

Where σ_{GSVI} is the standard deviation of the preceding 52 weeks.

The weekly GSVI downloaded is reported from Sunday to Sunday. However, the trading week for financial activity goes from Monday to Friday. As it is detailed in section 3.1.3 Financial variables, weekly stock returns are computed transforming daily closing prices into Monday-to-Monday week. Hence, the reported period of the GSVI is matched with the holding period of the given return, ensuring that both data sets are comparable and that the GSVI information is relevant for interpreting the return in question. Figure 9 illustrates the procedure for matching the GSVI and financial dataset.

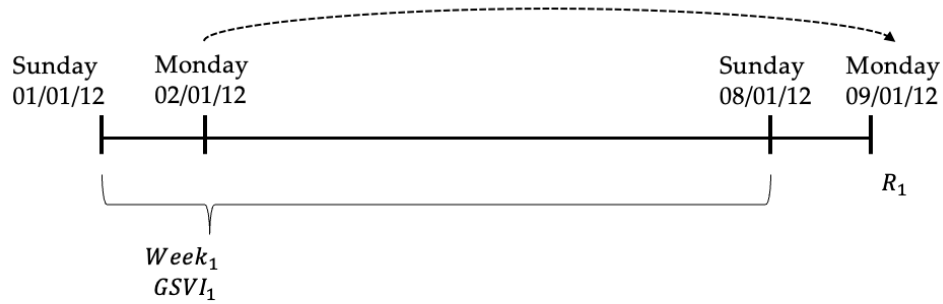


Figure 9. Representation of the matching process of GSVI and financial variables.

Source: Own Elaboration

5.1.3. Financial Variables

This section presents the financial variables used in the study's models. Given that multiple approaches are considered for calculating the variables, Table 5 outlines and describes the financial and search volume variables employed in the study, with the aim of ensuring greater clarity.

Table 5. List of variables and description.

Variable	Description	Source
ER_t	Weekly Excess return	Author calculation
AR_t	Weekly Abnormal return using Fama and French 3-factor model	Author calculation
$Volatility_t$	Weekly volatility	Molnár, P. (2012)
DTV_t	Detrended weekly trading volume	Campbell et al. (1993)
ATV_t	Abnormal weekly trading volume	Cheraghali et al. (2023)
$AGSVI_{DA8,t}$	Abnormal Google Search volume index	Da et al. (2011)
$AGSVI_{DA52,t}$	Abnormal Google Search volume index	Bijl et al. (2016)
$AGSVI_{BIJL52,t}$	Abnormal Google Search volume index	Bijl et al. (2016)

Source: own elaboration.

5.1.3.1. *Stock Return*

The first financial variable to estimate is the stock return. This metric measures the gain or loss generated by a stock investment over a certain period of time. Returns are useful for measuring the performance of different assets because, unlike prices, they create a comparative framework by normalizing the data. However, although stock return is considered an essential measure, advanced financial models require more sophisticated approaches that consider additional factors to provide a more accurate evaluation and prediction of market behaviour and individual assets.

The scientific community has proposed different approaches for this financial metric. The most commonly used ones in the field of this study are Abnormal return and Excess return. Both the Abnormal return and the Excess return are metrics widely used by the academia to measure the effect of the GSVI on financial markets. Therefore, this thesis incorporates both approaches in its methodology to assess stock returns. Below, the variables are detailed, along with a literature review and the calculation methodology for each one.

Firstly, Abnormal return measures the difference between the actual return of a security and the expected return based on an asset pricing model such as the Capital Asset Pricing Model (CAPM) (Sharpe, 1964; Lintner, 1965) or their extension as Fama and French (1993) three-factor model (Takeda and Wakao, 2014; Adachi et al., 2017; Kim et al., 2019), Carhart (1997) four-factor model (Tazi et al., 2022; Khan and Fahim, 2021; Artmann et al., 2012) and Fama and French (2015) five-factor model (Huang, 2019; Dirkx and Peter, 2020; Acaravci and Karaomer, 2017).

The literature on the GSVI calculates the Abnormal return using various extensions of the CAPM. El Ouadghiri et al. (2021), who study the effect of climate-related interest on the US market, extend the four-factor asset pricing model (Carhart, 1997) for estimating abnormal returns, by including other factors such as the GSVI for climate change and pollution interest. Following Joseph et al. (2011) and Takeda and Wakao (2014), Adachi et al. (2017) applied Fama and French (1993) three-factor model for Abnormal return in JASDAQ stock market. The authors argued that the momentum factor from the four-factor model is not appropriated for the Japanese market. Kim et al. (2019) used the Fama and French five-factor model to adjust the return variable in the Norwegian market.

To calculate the Abnormal Return, the Fama and French (1993) three-factor pricing model is regressed, and the result is subtracted from the actual weekly return.

The actual return is computed based on the equation below:

$$R_t = \frac{P_t - P_{t-1}}{P_{t-1}} \quad (6)$$

Where R_t is the weekly return, P_t is the adjusted close price in Monday of week t and P_{t-1} is the adjusted close price in Monday of week $t-1$. The same frequency as the weekly data from *Google Trends* is maintained, as it provides the average searches from Sunday to Sunday.

The expected return is computed following Fama and French (1993) three-factors model:

$$R - R_f = \alpha + \beta_m(R_m - R_f) + \beta_1SMB + \beta_2HML + \varepsilon \quad (7)$$

Where R is the expected return, R_f is the risk-free rate, R_m is the market return, SMB is the size factor (Small minus big) and HML is the value factor (high minus low). The risk-free rate and pricing factor SMB and HML are daily downloaded from Kenneth R. French's data library and transformed afterwards into weekly frequency by compounding the returns. Daily⁷ European Fama and French three-factors database is used.

The beta coefficients are estimated on a rolling window of 52-weeks with at least 25⁸ weeks of returns preceding week t .

Finally, the AR is calculated:

⁷ Weekly frequency is available only for US market.

⁸ This study established a minimum duration of 25 weeks to account for companies that were not continuously part of the index throughout the entire sample period. This approach ensures a sufficiently representative sample.

$$AR = R - R_f - \widehat{\beta}_m(R_m - R_f) + \widehat{\beta}_1SMB + \widehat{\beta}_2HML \quad (8)$$

Secondly, the Excess return represents the additional return that an investment earns compare to a benchmark or risk-free rate. It is used to assess the performance of an investment according to its level of risk and to optimize the risk-return profile. Evidence regarding the use of Excess return can be found in the works of Challet and Ayed (2013) and Preis et al. (2013), who measured the effect on the GSVI over market returns. Contrary, other authors as Cheraghali et al. (2023) Swamy et al. (2019) and Bijl et al. (2016) estimated Excess return for studying the specific-firm impact. Swamy et al. (2019) studied the Indian market using as benchmark the S&P BSE 500 index. Similarly, Bijl et al. (2016) measured the impact of the GSVI calculating the Excess return of individual companies listed in the S&P500 index by subtracting the 52-week rolling beta of each company multiply by the index return to individual company stock return. Cheraghali et al. (2023) also considered as dependent variable the excess return as the weekly return minus the risk-free rate of the components of S&P Global Luxury Index to measure consumer attention predictability.

Following Chronopoulos et al., (2018) and Cheraghali et al., (2023), Excess return is calculated by subtracting the risk-free rate to the weekly return. Firstly, daily stock prices are downloaded and then transformed into weekly stock return following the equation⁹ :

$$R_t = \frac{P_t - P_{t-1}}{P_{t-1}} \quad (9)$$

Where R_t is the weekly return, P_t is the adjusted close price in Monday of week t and P_{t-1} is the adjusted close price in Monday of week $t-1$. The study maintains the same frequency as the weekly data from *Google Trends*, as it provides the average searches from Monday to Sunday.

⁹ Dividends are not added in the formula because closing prices already include them.

A particular characteristic of this study is that EUROSTOXX50 includes companies listed on various European stock exchanges, each operating with different opening and closing days in accordance with the official calendar of their respective countries. In the event that Monday is a holiday, the adjusted closing price of the subsequent trading day is used.

For calculating excess return, the weekly risk-free rate is subtracted from the weekly return. In this manner, we measure the risk-free rate impact by weekly fluctuations. The risk-free rate is firstly daily downloaded from Kenneth R. French's data library¹⁰ and then converted into weekly data by compounding. The following equation presents the compound weekly risk-free rate, where n is the number of trading days in a week, and $R_{f,d}$ the daily risk-free rate.

$$R_{f,t} = \prod_{i=1}^n (1 + R_{f,d}) - 1 \quad (10)$$

Excess Return is presented as:

$$ER_{i,t} = R_{i,t} - R_{f,t} \quad (11)$$

5.1.3.2. *Volatility*

The second financial variable introduced is volatility. Volatility is a statistical measure that quantifies the extent to which the price of an asset fluctuates.

This variable estimation has been the subject of debate within the field of finance (Poon and Granger, 2003). Depending on the specific type of volatility being estimated, it may be regarded as either directly observable or not directly observable variable.

Researcher have examined the link between the GSVI and different measures of stock volatility: historical volatility (HV), realized volatility (RV), and implied

¹⁰ https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

volatility (IV). The primary difference among these measures lies in the data used for their calculation: Historical volatility relies on past stock prices; Realized volatility focuses on intraday prices, recently observed data; Implied volatility is centred on market expectations.

Table 6 shows a sample of publications that study the relation between the GSVI and volatility.

Table 6. Classification of Articles by Market and Volatility Approach

PUBLICATION	MARKET	VOLATILITY
Vlastakis and Markellos (2012)	DJIA	HV, RV and IV
Ramos et al., (2019)	EUROSTOXX50	HV
Dimpfl and Jank (2016)	DJIA	RV
Hamid and Heiden (2015)	DJIA	RV
Basistha et al., (2019)	S&P500 and DJIA	IV and RV
Chronopoulos et al., (2018)	S&P500	RV
Dzielinski M (2012)	S&P500	IV and RV
Aouadi et al., (2013)	CAC40	HV
González-Fernandez and González-Velasco (2020)	CAC40, DAX30, FTSE-MIB and IBEX35	IV

Source: Own elaboration.

Vlastakis and Markellos (2012), in their study over DJIA index, concluded that information demand, measured by the GSVI, has a significant and positive relation to both historical and realized volatility. Similarly, Dimpfl and Jank (2016) estimated the realized volatility computing intraday log-price changes over 10 minutes intervals of the same index. The authors concluded that including the Google searches in AR models improves volatility forecast.

Implied volatility is often assessed through volatility indices, typically from the VIX index family, which rely on the option prices of various benchmark stock indices. Batishta et al. (2019) measured implied volatility using de CBOE VIX index for the S&P 500 Index and DJIA Index options. González-Fernandez and González-Velasco (2020) measured investor sentiment to the COVID-19 crisis in Europe along with implied volatility via VIX index. Dzielinski, M (2012) measured the economic

uncertainty analysing the relation between implied volatility and Google search traffic.

Although it is considered as a less accurate tool because of the amount of data necessary for modelling it (Vlastakis and Markellos, 2012), the historical volatility is also a popular variable. In Molnár, P. (2012) words “high frequency data are in many cases not available at all or available only over a shorter time horizon and costly to obtain and work with” (p.20).

Following Molnár, P. (2012) study, we use Garman and Klass (1980) to compute realized volatility. The variable is calculated using daily closing ($close_d$), opening ($open_d$), high ($high_d$) and low (low_d) prices according to the equation below:

$$Variance_d = \frac{1}{2} (h_d - l_d)^2 - (2\log 2 - 1)c_d^2 + j_d^2 \quad (12)$$

Where

$$\begin{aligned} h_d &= \log(high_d) - \log(open_d) \\ l_d &= \log(low_d) - \log(open_d) \\ c_d &= \log(close_d) - \log(open_d) \\ j_d &= \log(open_d) - \log(close_{d-1}) \end{aligned} \quad (13)$$

The weekly volatility is calculated as the square root of the accumulated daily variance in week t .

$$VOL_t = \sqrt{\sum_{d \in t} \sigma_d^2} \quad (14)$$

5.1.3.3. *Trading volume*

The third financial variable that we construct is trading volume. The stock trading volume measures market activity indicating how much a security has been traded during a period of time. This tool is an indicator for liquidity (Adachi et al.,

2017; Darolles et al., 2015; Lo and Wang, 2010) and signal for market sentiment (Liu, S., 2015; So and Lei, 2015; Joseph et al., 2011).

Trading volume is computed following two different approaches: detrended trading volume and abnormal trading volume.

The detrended trading volume permits to isolate short-term fluctuations and cycles. Bijl et al., (2016) and Swamy et al., (2017) are followed, who calculate detrended trading volume (DTV) based on Campbell et al. (1993). The trading volume is the log volume in week t minus the rolling average of the past 12 weeks of log volume:

$$DTV_t = \log(\text{Volume}_t) - \frac{1}{12} \sum_{i=t-11}^t \log(\text{Volume}_i) \quad (15)$$

The abnormal trading volume is more popular among the literature that study the impact of investor attention in financial activity (Lai et al., 2022; Nguyen et al., 2020; Padungsaksawasdi, C., 2019; Adachi et al., 2019). Abnormal trading volume allows to identify unusual events such as release of new information or changes in market sentiment and investor attention.

Abnormal trading volume denoted as ATV is estimated following Cheraghali et al. (2023). Firstly, it is calculated weekly trading volume from daily data.

$$ATV_t = \frac{1}{S} \sum_{i=1}^S TV_i \quad (16)$$

Where S is the number of trading days in week t .

Then, abnormal volume is calculated as:

$$ATV_t = \frac{TV_t - \frac{1}{52} \sum_{i=1}^{52} TV_{t-i}}{\sigma_{TV}} \quad (17)$$

Where ATV_t is the abnormal trading volume in week t and σ_{TV} is the standard deviation of trading volume for the year preceding week t .

5.2. RESEARCH MODELS

To study the impact of the AGSVI over the market activity I apply panel data regression models to analyse the current and future relation between the AGSVI and the financial metrics. From this point onward, in the context of the model definitions, the term “return” will encompass both abnormal and excess returns, while “trading volume” will refer to both abnormal trading volume and detrended trading volume, with the aim of providing greater clarity for the reader.

5.2.1. Panel data.

Panel data, also known as longitudinal or cross-sectional time-series data, consists of observations on multiple entities (such as firms, countries, or individuals) over time. In finance, panel data typically involves variables measured across different financial assets (e.g., stocks) or firms over a specified time period. In this thesis, a dataset of stock returns, volatility, trading volume and the AGSVI across multiple companies over several years would be considered panel data.

Panel data combines both cross-sectional and time-series dimensions, allowing researchers to track changes over time for each entity while also examining differences between entities at specific points in time. This dual nature offers several advantages for econometric analysis, making panel data a robust choice for studying financial and attention variables (Brooks, 2019).

In econometrics, it is identified as follows:

$$y_{i,t} = \alpha + \beta_1 x_{i,t} + \mu_i \quad (18)$$

Where $y_{i,t}$ is the dependent variable, α is the intercept term, $x_{i,t}$ is a $k \times 1$ vector of parameters to be estimated on the explanatory variables, and is a $k \times 1$ vector of observations on the explanatory variables, $t=1,\dots,T$; $i=1,\dots,N$.

¹¹ Note that k corresponds to the number of independent variables in the regression model.

Panel data captures both the variation across entities and over time, leading to greater informational richness. This allows for more nuanced insights into how financial variables interact. For instance, the relationship between stock returns and the AGSVI can be studied not only across different firms, using the 52 stock tickers, but also how that relationship evolves over time. In this thesis study, 10 years of 52 weeks observation per year.

The literature that studies the relation between the AGSVI and financial market behaviour has extensively utilized panel data models. For instance, Kim et al. (2019) employed this methodology on a sample of 28 companies from the Norwegian market over a five-year period with weekly frequency. Similarly, Cheraghali et al. (2023) utilized panel data to examine the components of the S&P Global Luxury Index over an 18-year period, adopting an analogous approach. Additional studies that have also employed panel data include Bijl et al. (2016), Yoshinaga and Rocco (2020), and Adachi et al. (2017). Given its widespread application within the field, this methodology is considered highly suitable for the present thesis.

A key point to be addressed is how to control for unobserved variables that could influence the relationship between the variables. For instance, firm-specific factors, such as the sector where the firm operates, that can affect $y_{i,t}$ cross-sectionally but remain constant through time (Brooks, 2019).

When analysing panel data that includes companies from the European market, it is assumed that a fixed effects model may be more appropriate than a random effects model. This is because European firms likely differ in ways that are constant over time but unobserved in the data, such as country-specific regulatory environments, management practices, corporate culture, or market conditions. These unobserved characteristics could influence key financial variables, like returns or volatility. A fixed effects model allows to control for these unobserved, time-invariant factors, ensuring they do not bias the results (Kim et al., 2019). In contrast, a random effects model assumes that these unobserved characteristics are uncorrelated with the independent variables, which may be unrealistic in this context, potentially leading to biased estimates. Furthermore, the literature on the use of the AGSVI in financial market studies predominantly employs fixed effects models in their methodology. Notable studies that have utilized fixed effects

models include Cheraghali et al. (2023); Yoshinaga and Rocco (2020); Kim et al. (2019); Adachi et al. (2017) and Bijl et al. (2016).

Therefore, given the nature of the data, the inherent differences across European companies, and the support from the literature, fixed effects are considered to be a more appropriate choice.

However, this analysis conducts the Hausman test (Hausmann, 1978) to have a more econometrically supported assumption (Bijl et al. 2016; Adachi et al. 2017). The central question the test addresses is whether the unobserved individual-specific effects (such as firm-specific characteristics in a dataset of companies) are correlated with the explanatory variables (Amini et al., 2012).

The Hausman test is conducted over the 162 regressions, the significant level is 0.05 and the following two hypothesis are proposed:

H₀: No correlation between the individual-specific effects and the regressors (random effects is appropriate).

H₁: Correlation exists between the individual-specific effects and the regressors (fixed effects is appropriate).

The results are presented in figure 10. Indeed, 122 over the 162 regressions yield a p-value lower than 0,05 significant level. Hence, the null hypothesis is rejected, confirming fixed effect model as the appropriated one.

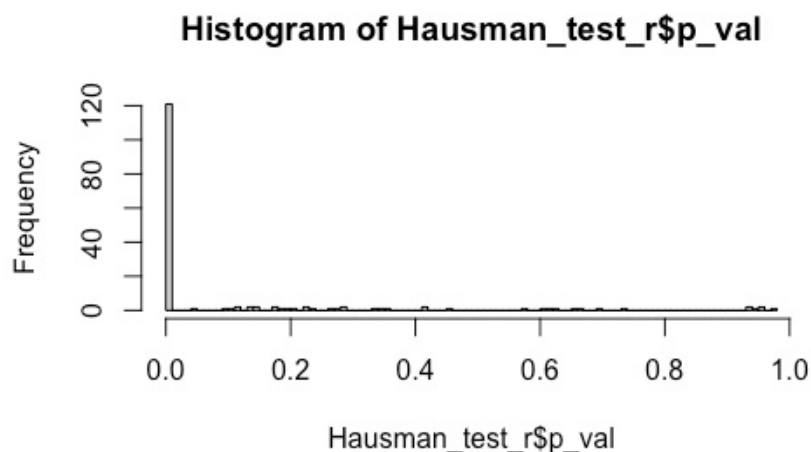


Figure 10. Hausman test.

Source: Own elaboration.

5.2.2. Cross-sectional and Predictive models.

The relation between the AGSVI and the financial variables is measured based on three different panel data regressions models. Firstly, I study the impact of investor attention over financial variables regressing them just against the AGSVI. The current relation between the variables is studied under a nowcasting approach (Smales, 2021; Ding et al., 2020; Nguyen et al., 2020; El Ouadguiri and Peillex, 2018; Aouadi et al., 2013; Bank et al., 2011). According to Choi and Varian, (2012), the GSVI can help in current forecasting. In other words, how past search intensity volume forecast present market activity.

The search intensity variable is also lagged in order to capture the forecasting capacity of the AGSVI. In this sense, it can be measured if past search queries data can predict future stock market activity. (Cheraghali et al., 2023; Salisu and Vo, 2021; Yoshinaga and Rocco, 2020; Swamy et al., 2019; Ramos et al., 2019; Chen, 2017; Moussa et al., 2017a,b; Dimpfl and Jank, 2016; Joseph et al., 2011).

The equation is as follow:

$$Y_{i,t} = \alpha + \beta_1 AGSVI_{i,t-n} + \mu_i + \varepsilon_{i,t} \quad (19)$$

Where β_1 is the regression coefficient, $Y_{i,t}$ is the dependent variable (return, volatility and trading volume) of company i at time t , μ_i is the fix effect by company. The AGSVI is either $AGSVI_{DA8,t}$; $AGSVI_{DA52,t}$; $AGSVI_{BIJL52,t}$ that proxies the investor attention, lagged n periods.

The second model extends Model (1) by adding the lagged dependent financial variable. Model (2) includes the lagged financial variable because there exist a dependency pattern and market memory. Financial markets often show patterns where the current value of a variable (like a stock price or return) is dependent on its past values. This dependency is a key feature of time series data where trends and patterns tend to persist for some time (Brooks, 2019). Hence, leveraging the information contained in past data, it can be improved the predictive power of the model and to understand the underlying processes driving financial markets. This idea is supported by current literature such as Chiang et al., (2010), Ané and Urenche-Rangau (2008) and Chuang et al., (2012), who demonstrated the utility of lagged values to predict return volatility.

The search intensity variable is also lagged to study its predictive capacity. Model (2) is displayed as follow:

$$Y_{i,t} = \alpha + \beta_1 Y_{i,t-1} + \beta_2 AGSVI_{i,t-n} + \mu_i + \varepsilon_{i,t} \quad (20)$$

Where, $Y_{i,t}$ is the dependent variable (return, volatility and trading volume) of company i at time t , $Y_{i,t-1}$ is the 1-week lagged dependent variable, β_1 and β_2 are the regression coefficients, μ_i is the fix effect by company. The AGSVI is either $AGSVI_{DA8,t}$; $AGSVI_{DA52,t}$; $AGSVI_{BIJL52,t}$ that proxies the investor attention, lagged n periods.

The third model extends the second by adding control variables to the regressions. The objective is to estimate investor attention and its impact on market variability, specifically in the short term. The control variables are one-week lagged and the AGSVI one and three weeks. Tantaopas et al. (2016) suggest that the optimal lag length for analysing return and volatility is three weeks. Furthermore, this thesis hypothesizes that when an investor seeks information about a company, his intention is to trade their assets within a brief timeframe. Hence, the model tests the predictive capacity first with three weeks lag and then one week-lag following literature. Model (3) is as follow:

$$Y_{i,t} = \alpha + \beta_1 Y_{i,t-1} + \beta_2 AGSVI_{i,t-n} + X' \gamma_{i,t-1} + \varepsilon_{i,t} \quad (21)$$

Where, β_1 and β_2 are the regression coefficients, $Y_{i,t}$ is the dependent variable (return, volatility and trading volume) of company i at time t , $Y_{i,t-1}$ are the 1-week lagged dependent variable and μ_i is the fix effect by company. X' is the control variables vector which are the financial metrics that are not used as dependent variable in each equation. For instance, when estimating $Y_{t,i} = AR_{t,i}$, the control variables are AR_{t-1} , Vol_{t-1} and TV_{t-1} . The AGSVI is either $AGSVI_{DA8,t}$; $AGSVI_{DA52,t}$; $AGSVI_{BIJL52,t}$ that proxies the investor attention, being $n = 1$ or 3 lags.

VI – EMPIRICAL RESULTS

VI -EMPIRICAL RESULTS

6.1. SUMMARY STATISTICS

This section presents the summary statistics of the variables used in the study. This analysis provides a comprehensive overview of the fundamental characteristics of the data, identifying potential patterns or anomalies that could influence the outcomes of the subsequent analysis.

Firstly, Table 7 exposes the descriptive statistics for all the variables. The abnormal return (AR_t) is calculated with Fama and French 3-factor model (Fama and French, 1993), the excess return (ER_t) by subtracting the risk-free rate to the weekly return. The realized volatility ($volatility_t$) is calculated based on Garman and Klass estimator, following Molnár, P., (2012). Trading volume variables correspond to abnormal trading volume (ATV_t) and detrended trading volume (DTV_t). Finally, the search intensity variables ($AGSVI_{DA8,t}$, $AGSVI_{DA52,t}$ and $AGSVI_{BIJL52,t}$) are estimated as specified in section 3.1.2.2.

The stocks return of the sample companies show significant variation in both abnormal and excess return, ranging from -69,6% to +76,1% and -63,6% to +67,6%, respectively. Similar results are obtained with the Google variable, being even more prominent the variation in $AGSVI_{BIJL52,t}$, ranging from -7.776% to 168.860%.

The descriptive data show significant outliers in some of the variables, which may be due to exceptional phenomena. In the data sample, for instance, in the return variable, some companies display more aggressive variations than usual due to an intra-week increase in the stock price. The nature of the sample constituents could naturally cause wide ranges between the maximum and minimum values. Although the EUROSTOXX50 index represents relatively stable companies in the context of the European market, differences exist, such as the sector to which they belong (Vo et al., 2023) or their stability over time.

Table 7. Descriptive statistics for all variables

Variable	N	Mean	St. Dev.	Min	Max
AR_t	27,481	0.002	0.046	-0.696	0.761
ER_t	27,481	0.001	0.042	-0.636	0.676
Volatility_t	27,476	0.037	0.023	0.004	1.012
ATV_t	27,346	0.033	4.355	-3.571	688.072
DTV_t	27,481	0.001	0.325	-2.285	5.780
AGSVI_{DA8,t}	27,481	0.008	0.166	-1.179	2.996
AGSVI_{DA52,t}	27,481	0.023	0.202	-1.299	3.219
AGSVI_{BIJL52,t}	27,481	0.107	1.719	-7.776	168.860

Source: Own elaboration.

Table 8 reports the correlation matrix of the proposed variables of the study. The results reveal low values of correlation excepts among the search intensity variables (AGSVI_{DA8,t} and AGSVI_{DA52,t}; AGSVI_{BIJL52,t}). This could implicate that the index proposed for investor attention can potentially provide relevant and explanatory information such as trends and trading sentiment for stock movements measured by the financial variables.

The AGSVI variables present positive correlations with all financial variables, with the exception of return (both abnormal and excess) where the relationship is negative. The correlation between the three measures of investor attention ranges from 0.6 to 0.7, indicating a certain degree of interdependence among them.

Table 8. Correlation Matrix for all the variables

	AR_t	ER_t	$volatility_t$	ATV_t	DTV_t	$AGSVI_{DA8,t}$	$AGSVI_{DA52,t}$	$AGSVI_{BIJL52,t}$
AR_t	1							
ER_t	0.936	1						
$volatility_t$	-0.166	-0.184	1					
ATV_t	-0.048	-0.056	0.138	1				
DTV_t	-0.140	-0.164	0.390	0.333	1			
$AGSVI_{DA8,t}$	-0.015	-0.010	0.044	0.009	0.068	1		
$AGSVI_{DA52,t}$	-0.018	-0.013	0.073	0.014	0.052	0.720	1	
$AGSVI_{BIJL52,t}$	-0.013	-0.008	0.038	0.009	0.037	0.604	0.672	1

Source: Own elaboration.

6.2. PANEL DATA REGRESSIONS

To assess the relationship between investor attention and market behaviour, this thesis is focused on three financial variables: return, volatility, and trading volume. This section presents the results of the three proposed regression models, measuring both the contemporaneous and predictive relationships between the GSVI variable and each of the three financial variables individually.

The hypotheses regarding the relationship between search intensity and each financial variable are tested using three different models. First, the GSVI is regressed on each financial variable. Second, the initial model is extended by incorporating the lagged dependent variable. Finally, the financial variable is analysed through a multivariate regression that includes additional financial variables. The proxy variable for investor attention is measured at t_0 to assess its contemporaneous relationship with returns and with lags of 1 and 3 weeks to evaluate its predictive ability. This analytical structure is replicated across the three pairs of hypotheses.

Given the extensive set of regressions to be estimated and the wide range of results obtained, each subsection of this chapter begins with a summary table outlining the significance, magnitude, and direction of the GSVI coefficients in relation to the exogenous financial variables. This approach aims to distill the key findings in a clear and effective manner. Subsequently, the results are presented in

greater detail through comprehensive tables, considering the number of observations and R^2 values.

6.2.1. GSVI and Return

This section provides an analysis of the results obtained from measuring the relationship between investor attention and stock returns. Thus, the hypotheses RH_{1a} and RH_{1b} are examined. Hypothesis RH_{1a} posits that an increase in search volume is linked to lower stock returns. In the same line, the hypothesis RH_{1b} is formulated, lagging the investor attention variable to test its predictive capacity over stock returns.

Firstly, for a clearer understanding of the results, the summary Table 9 reports the coefficients of the primary variable under study, AGSVI, without presenting the values of the additional variables included in the multiple regression models. The subsequent tables will provide a detailed analysis of the effects of these additional variables on the dependent variable.

Table 9. Summary results on AR and ER

		Abnormal Return			Excess Return		
		t0	t-1	t-3	t0	t-1	t-3
Model (1)	AGSVI _{DA8,t}	-0.004	-0.003*	-0.003	-0.003	-0.002	-0.002
	AGSVI _{DA52,t}	-0.005*	-0.003**	-0.002	-0.003	-0.002	-0.001
	AGSVI _{BIJL52,t}	-0.0004	-0.0002	-0.0004***	-0.0002	-0.0001	-0.0003**
Model (2)	AGSVI _{DA8,t}	-0.004	-0.003*	-0.003	-0.003	-0.002	-0.002
	AGSVI _{DA52,t}	-0.005*	-0.003**	-0.002	-0.003	-0.002	-0.001
	AGSVI _{BIJL52,t}	-0.0004	-0.0002	-0.0004***	-0.0002	-0.0001	-0.0003**
Model (3)	ATV _t	AGSVI _{DA8,t}	-0.001	-0.001	-0.001	-0.002	
		AGSVI _{DA52,t}	-0.001	-0.001	-0.001	-0.0004	
		AGSVI _{BIJL52,t}	-0.00005	-0.0003**	-0.00001	-0.0003**	
	DTV _t	AGSVI _{DA8,t}	-0.001	-0.002	-0.001	-0.002	
		AGSVI _{DA52,t}	-0.001	-0.001	-0.001	-0.001	

AGSVI _{BIJL52,t}	-0.00004	-0.0003**	-0.00000	-0.0003**
---------------------------	----------	-----------	----------	-----------

Note: Models (1) regresses the AGSVI against stock Return. Model (2) builds upon Model (1) by incorporating the lagged Return variable. Model (3) extends Model (2) by including lagged volatility and trading volume. The sample period covers weekly data from 2014 to 2024. The symbols ***, ** and * denote significance at the 1%, 5% and 10% levels, respectively.

Source: Own Elaboration.

To initiate the analysis, it is conducted Model (1), where the return variable is regressed on the AGSVI (AGSVI_{DA8} ; AGSVI_{DA52} and AGSVI_{BIJL52})¹² which serves as the sole independent variable in the model. Both contemporaneous and predictive relationships are examined, with the search volume variable lagged by 1 and 3 weeks. The results of these regressions are reported in Tables 10 and 11. The sample period spans from 2014 to 2024, with weekly frequency data across a total of 52 firms. As a result, the full sample includes approximately 27,481 observations when the return and AGSVI variables align within the same period, and 27,427 (27,320) observations when the investor attention variable is lagged by 1 (3) weeks.

Overall, the results indicate a negative relationship between the AGSVI and return. However, the relation is barely significant when return is measured using the excess return variable. The analysis reveals that an increase of one unit in Google searches is associated with a decrease in return, but only when searches are measured using the AGSVI_{DA52} variable. As the investor attention variable is lagged, the significance level of the results increases. An increase in investor attention at $t-3$ is associated with a decrease of 0.0004 points in return at t , with a 1% significance level.

Although the values for R^2 of the three models may be relatively small, similar levels have been reported in prior studies, including Cheraghali et al. (2023); Swamy et al. (2019); Kim et al. (2019) and Bijl et al. (2016). Moreover, the values increase as Model (1) is expanded with additional explanatory variables.

¹² From this point onward, when referring to AGSVI, it will denote the variable in its three specification forms: AGSVI_{DA8,t} , AGSVI_{DA52,t} and AGSVI_{BIJL52,t}

Table 10. Regression results on Abnormal Return with the AGSVI as independent variable.

	Abnormal Return								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
AGSVI _{DA8,t}	-0.004 (0.003)								
AGSVI _{DA52,t}		-0.005* (0.003)							
AGSVI _{BIJL52,t}			-0.0004 (0.0004)						
AGSVI _{DA8,t-1}				-0.003* (0.002)					
AGSVI _{DA52,t-1}					-0.003** (0.002)				
AGSVI _{BIJL52,t-1}						-0.0002 (0.0002)			
AGSVI _{DA8,t-3}							-0.003 (0.002)		
AGSVI _{DA52,t-3}								-0.002 (0.001)	
AGSVI _{BIJL52,t-3}									-0.0004*** (0.0001)
Observations	27,481	27,481	27,481	27,427	27,427	27,427	27,320	27,320	27,320
R ²	0.0002	0.0004	0.0002	0.0001	0.0002	0.00004	0.0001	0.0001	0.0002

Note: Models (1) to (3) assess the current relation of AGSVI over Abnormal Return. Columns (4) to (6) display results of forecasting model of AGSVI 1-week lagged. Columns (7) to (9) display results of forecasting model of AGSVI 3-week lagged. Robust standard errors are clustered by ticker and time are reported in parentheses. The sample period covers weekly data from 2014 to 2024. The symbols ***, ** and * denote significance at the 1%, 5% and 10% levels, respectively.

Source: Own Elaboration.

Table 11. Regression results on Excess Return with the AGSVI as independent variable.

	Excess Return								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
AGSVI _{DA8,t}	-0.003 (0.003)								
AGSVI _{DA52,t}		-0.003 (0.002)							
AGSVI _{BIJL52,t}			-0.0002 (0.0003)						
AGSVI _{DA8,t-1}				-0.002 (0.001)					
AGSVI _{DA52,t-1}					-0.002 (0.001)				
AGSVI _{BIJL52,t-1}						-0.0001 (0.0002)			
AGSVI _{DA8,t-3}							-0.002 (0.002)		
AGSVI _{DA52,t-3}								-0.001 (0.001)	
AGSVI _{BIJL52,t-3}									-0.0003** (0.0001)
Observations	27,481	27,481	27,481	27,427	27,427	27,427	27,320	27,320	27,320
R ²	0.0001	0.0002	0.0001	0.00005	0.0001	0.00001	0.0001	0.00004	0.0001

Note: Models (1) to (3) assess the current relation of AGSVI over Excess Return. Columns (4) to (6) display results of forecasting model of AGSVI 1-week lagged. Columns (7) to (9) display results of forecasting model of AGSVI 3-week lagged. Robust standard errors are clustered by ticker and time are reported in parentheses. The sample period covers weekly data from 2014 to 2024. The symbols ***, ** and * denote significance at the 1%, 5% and 10% levels, respectively.

Source: Own Elaboration.

The results are similar when testing Model (2). Tables 12 and 13 present the findings of the panel data with return as dependent variable and the AGSVI and 1-week lagged return as explanatory variables. This model examines whether Google searches have an additional impact on the current return, controlling for the effect

of the previous week return. Similar to the first model, the results are more significant for abnormal return than for excess return.

In the autoregressive model of abnormal return, the results show that a high return in the previous period is associated with a decrease of 0.061 points in the current return, with a significance level of 1%. Therefore, when the return increases, there tends to be a decline in the following period's return. This effect is not significant when the relationship is estimated using excess return. Regarding the effect of investor attention, the same findings as in Model (1) are observed. Specifically, an increase in search intensity three weeks prior is negatively related to current week returns, although these results only emerge when the abnormal GSVI is estimated using the AGSVI_{BIJL52}.

Table 12. Regression results on Abnormal Return with the AGSVI and lagged AR as independent variable.

	Abnormal Return								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
AR _{t-1}	-0.061*** (0.010)	-0.061*** (0.010)	-0.061*** (0.010)	-0.061*** (0.010)	-0.061*** (0.010)	-0.061*** (0.010)	-0.062*** (0.010)	-0.062*** (0.010)	-0.062*** (0.010)
AGSVI _{DA8,t}	-0.004 (0.003)								
AGSVI _{DA52,t}		-0.005* (0.003)							
AGSVI _{BIJL52,t}			-0.0004 (0.0004)						
AGSVI _{DA8,t-1}				-0.003* (0.002)					
AGSVI _{DA52,t-1}					-0.003** (0.002)				
AGSVI _{BIJL52,t-1}						-0.0002 (0.0002)			
AGSVI _{DA8,t-3}							-0.003 (0.002)		
AGSVI _{DA52,t-3}								-0.002 (0.001)	
AGSVI _{BIJL52,t-3}									-0.0004*** (0.0001)
Observations	27,427	27,427	27,427	27,427	27,427	27,427	27,320	27,320	27,320

R ²	0.004	0.004	0.004	0.004	0.004	0.004	0.004	0.004	0.004
----------------	-------	-------	-------	-------	-------	-------	-------	-------	-------

Note: Models (1) to (3) assess the current relation of AGSVI and lagged abnormal return over abnormal return. Columns (4) to (6) display results of forecasting model of the AGSVI 1-week lagged. Columns (7) to (9) display results of forecasting model of the AGSVI 3-week lagged. Robust standard errors are clustered by ticker and time are reported in parentheses. The sample period covers weekly data from 2014 to 2024. The symbols ***, ** and * denote significance at the 1%, 5% and 10% levels, respectively.

Source: Own Elaboration.

Table 13. Regression results on excess return with the AGSVI and lagged excess return as independent variable

	Excess Return								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
ER _{t-1}	-0.006 (0.011)	-0.007 (0.011)	-0.006 (0.011)	-0.006 (0.011)	-0.007 (0.011)	-0.006 (0.011)	-0.007 (0.011)	-0.007 (0.011)	-0.007 (0.011)
AGSVI _{DA8,t}	-0.003 (0.003)								
AGSVI _{DA52,t}		-0.003 (0.002)							
AGSVI _{BIJL52,t}			-0.0002 (0.0003)						
AGSVI _{DA8,t-1}				-0.002 (0.001)					
AGSVI _{DA52,t-1}					-0.002 (0.001)				
AGSVI _{BIJL52,t-1}						-0.0001 (0.0002)			
AGSVI _{DA8,t-3}							-0.002 (0.002)		
AGSVI _{DA52,t-3}								-0.001 (0.001)	
AGSVI _{BIJL52,t-3}									-0.0003** (0.0001)
Observations	27,427	27,427	27,427	27,427	27,427	27,427	27,320	27,320	27,320
R ²	0.0002	0.0003	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0002

Note: Models (1) to (3) assess the current relation of AGSVI and lagged Excess return over the dependent variable. Columns (4) to (6) display results of forecasting model of AGSVI 1-week lagged. Columns (7) to (9) display results of forecasting model of AGSVI 3-week

lagged. Robust standard errors are clustered by ticker and time are reported in parentheses. The sample period covers weekly data from 2014 to 2024. The symbols ***, ** and * denote significance at the 1%, 5% and 10% levels, respectively.

Source: Own Elaboration.

Finally, Tables 14 and 15 present the results of Model (3) with respect to the return variable. This model extends the autoregressive Model (2) by incorporating volatility and trading volume as control variables. The findings for the abnormal return indicate that the volatility from the previous period accounts for a significant portion of the variation in returns, with a 1% level of significance. An increase in volatility is associated with a subsequent decline in returns. Similar results are observed for trading volume, though the effect is smaller (-0.005 points of return). This could be attributed to negative investor expectations or economic uncertainty. The effect of the control variables remains consistent when the outcome variable is Excess Return.

Regarding the study's primary variable (AGSVI), the significance of investor attention in relation to returns disappears, except when the AGSVI is estimated by AGSVI_{BIJL52} variable lagged three weeks. An increase in investor attention at t-3 is associated with lower returns in the current period. This confirms that the GSVI adds value in explaining prices fluctuations, as it captures a portion of return behaviour that is not explained by other control variables such as lagged volatility and trading volume. These findings are consistent with the results of in Model (1) and (2), confirming that AGSVI_{BIJL52} is a robust predictor of stock returns.

After analysing the previous models, the following conclusions based on the research hypothesis are drawn:

Hypothesis RH_{1a} proposed that the GSVI is positively associated with stock returns. However, the results indicate that an increase of one unit in attention demand is linked to a decline in stock returns. This relationship is observed only when investor attention is estimated using the AGSVI_{DA52} variable, with a significance level of 10%. Given that the findings suggest a negative relationship between investor attention and stock returns, and this result is statistically significant in only one out of the three AGSVI estimation methods, RH_{1a} is rejected.

Hypothesis RH_{1b} states that the GSVI predicts stock returns. The results from Models (1) and (2) indicate that the 1-week lagged AGSVI variable is significantly associated with negative abnormal stock returns. However, once control variables are introduced in Model (3), this effect disappears. On the other hand, the most consistent and significant results occur three weeks after a change (increase) in Google search volume. An increase in investor attention at $t-3$, estimated by $AGSVI_{BIJL52}$, is associated with lower abnormal and excess returns at significance levels of 1% and 5%, respectively. Hence, it is considered that high levels of GSVI can predict lower returns. This 3-week lagged result in the GSVI holds in all model specifications, irrespective of the added control variables. However, although this pattern of results is consistent, it only occurs in one of the three approaches to estimating GSVI. Based on the obtained results, hypothesis RH_{1b} is partially accepted.

Table 14. Regression results on abnormal return with various independent variables

	Abnormal Return											
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
AR_{t-1}	-0.084*** (0.008)	-0.084*** (0.008)	-0.084*** (0.008)	-0.085*** (0.008)	-0.085*** (0.008)	-0.085*** (0.008)	-0.084*** (0.008)	-0.084*** (0.008)	-0.084*** (0.008)	-0.086*** (0.008)	-0.086*** (0.008)	-0.086*** (0.008)
$AGSVI_{DA8,t-1}$	-0.001 (0.001)			-0.001 (0.001)								
$AGSVI_{DA52,t-1}$		-0.001 (0.001)			-0.001 (0.001)							
$AGSVI_{BIJL52,t-1}$			-0.00005 (0.0002)			-0.00004 (0.0001)						
$AGSVI_{DA8,t-3}$							-0.002 (0.002)			-0.002 (0.002)		
$AGSVI_{DA52,t-3}$								-0.001 (0.001)			-0.001 (0.001)	
$AGSVI_{BIJL52,t-3}$									-0.0003** (0.0001)			-0.0003** (0.0001)
$Volatility_{t-1}$	-0.258*** (0.035)	-0.257*** (0.035)	-0.258*** (0.035)	-0.231*** (0.037)	-0.231*** (0.037)	-0.231*** (0.037)	-0.257*** (0.035)	-0.257*** (0.035)	-0.257*** (0.035)	-0.230*** (0.037)	-0.230*** (0.037)	-0.230*** (0.038)
ATV_{t-1}	-0.0002 (0.0002)	-0.0002 (0.0002)	-0.0002 (0.0002)				-0.0002 (0.0002)	-0.0002 (0.0002)	-0.0002 (0.0002)			
DTV_{t-1}				-0.005*** (0.001)	-0.005*** (0.001)	-0.005*** (0.001)				-0.005*** (0.001)	-0.005*** (0.001)	-0.005*** (0.001)
Observations	27,287	27,287	27,287	27,422	27,422	27,422	27,190	27,190	27,190	27,315	27,315	27,315
R ²	0.019	0.019	0.019	0.020	0.020	0.020	0.019	0.019	0.019	0.020	0.020	0.020

Models (1) to (6) report the regression results of the 1-week lagged financial metrics and the AGSVI to forecast Abnormal return. Models (7) to (12) report the regression results of the 1-week lagged financial metrics and the 3-weeks lagged AGSVI to forecast Abnormal return. Robust standard errors are clustered by ticker and time are reported in parentheses. The sample period covers weekly data from 2014 to 2024. The symbols ***, ** and * denote significance at the 1%, 5% and 10% levels, respectively.

Source: Own elaboration.

Table 15. Regression results on Excess return with various independent variables.

	Excess Return											
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
AR_{t-1}	-0.021** (0.009)	-0.021** (0.009)	-0.021** (0.009)	-0.024*** (0.009)	-0.024*** (0.009)	-0.024*** (0.009)	-0.021** (0.009)	-0.021** (0.009)	-0.021** (0.009)	-0.024*** (0.009)	-0.024*** (0.009)	-0.024*** (0.009)
$AGSVI_{DA8,t-1}$	-0.001 (0.001)			-0.001 (0.001)								
$AGSVI_{DA52,t-1}$		-0.001 (0.001)			-0.001 (0.001)							
$AGSVI_{BIJL52,t-1}$			-0.00001 (0.0001)			-0.00000 (0.0001)						
$AGSVI_{DA8,t-3}$							-0.002 (0.002)			-0.002 (0.002)		
$AGSVI_{DA52,t-3}$								-0.0004 (0.001)			-0.001 (0.001)	
$AGSVI_{BIJL52,t-3}$									-0.0003** (0.0001)			-0.0003** (0.0001)
$Volatility_{t-1}$	-0.134*** (0.028)	-0.134*** (0.028)	-0.135*** (0.028)	-0.111*** (0.029)	-0.110*** (0.029)	-0.111*** (0.029)	-0.134*** (0.028)	-0.134*** (0.028)	-0.134*** (0.028)	-0.110*** (0.029)	-0.110*** (0.030)	-0.110*** (0.030)
ATV_{t-1}	-0.0002 (0.0002)	-0.0002 (0.0002)	-0.0002 (0.0002)				-0.0002 (0.0002)	-0.0002 (0.0002)	-0.0002 (0.0002)			
DTV_{t-1}				-0.005*** (0.001)	-0.005*** (0.001)	-0.005*** (0.001)				-0.005*** (0.001)	-0.005*** (0.001)	-0.005*** (0.001)
Observations	27,287	27,287	27,287	27,422	27,422	27,422	27,190	27,190	27,190	27,315	27,315	27,315
R ²	0.003	0.003	0.003	0.004	0.004	0.004	0.003	0.003	0.003	0.004	0.004	0.004

Models (1) to (6) report the regression results of the 1-week lagged financial metrics and the AGSVI to forecast Excess return. Models (7) to (12) report the regression results of the 1-week lagged financial metrics and the 3-weeks lagged AGSVI to forecast Excess return. Robust standard errors are clustered by ticker and time are reported in parentheses. The sample period covers weekly data from 2014 to 2024. The symbols ***, ** and * denote significance at the 1%, 5% and 10% levels, respectively.

Source: Own elaboration.

6.2.2. GSVI and Volatility

After examining the impact of investor attention on return, this section explores its effect on volatility. The hypotheses RH_{2a} posits that an increase in investor attention is related to high volatility, as investors respond with more information-seeking in uncertainty situations. Moreover, the GSVI is delayed to test is predictive ability of volatility in European markets, as stated in hypothesis RH_{2b} .

As in previous section, it is firstly displayed a summary of the results with volatility as the dependent variable. Table 16 focuses exclusively on the coefficients of the key variable under study, AGSVI. The values of the additional variables included in the multiple regression model are not presented in this table, as their explanatory effect on the dependent variable will be analysed in detail in the following sections.

Table 16. Summary results on volatility

		Volatility		
		t0	t-1	t-3
Model (1)	AGSVI _{DA8,t}	0.006	0.003	0.003**
	AGSVI _{DA52,t}	0.008**	0.006***	0.005***
	AGSVI _{BIJL52,t}	0.001	0.0003	0.0003**
Model (2)	AGSVI _{DA8,t}	0.0003	-0.0003	0.002***
	AGSVI _{DA52,t}	0.002	0.002**	0.003***
	AGSVI _{BIJL52,t}	0.0001	-0.00000	0.0002**
Model (3)	AR_{t-1}		-0.0004	0.002***
	$+ ATV_{t-1}$			
	AGSVI _{DA8,t}		0.001**	0.003***
	AGSVI _{DA52,t}		-0.00001	0.0002**
	AGSVI _{BIJL52,t}			
	AR_{t-1}		-0.0004	0.002***
	$+ DTV_{t-1}$			
	AGSVI _{DA8,t}		0.002**	0.003***
	AGSVI _{DA52,t}		-0.00001	0.0002**
	AGSVI _{BIJL52,t}			

ER_{t-1} + ATV_{t-1}	$AGSVI_{DA8,t}$	-0.0003	0.002**
	$AGSVI_{DA52,t}$	0.002**	0.003***
	$AGSVI_{BIJL52,t}$	-0.00000	0.0002**
ER_{t-1} + DTV_{t-1}	$AGSVI_{DA8,t}$	-0.003	0.002***
	$AGSVI_{DA52,t}$	0.002**	0.003***
	$AGSVI_{BIJL52,t}$	-0.00000	0.0002**

Note: Models (1) regresses the AGSVI against Volatility. Model (2) builds upon Model (1) by incorporating the lagged volatility variable. Model (3) extends Model (2) by including lagged return and trading volume. The sample period covers weekly data from 2014 to 2024. The symbols ***, ** and * denote significance at the 1%, 5% and 10% levels, respectively.

Source: Own elaboration.

First, Model (1) assesses the effect on volatility while isolating the investor attention variable. For this purpose, volatility is regressed on the AGSVI, lagged by 1 and 3 weeks. Table 17 presents the regression coefficients for the three different measures of investor attention and their impact on weekly stock volatility. Overall, a positive relationship is observed between Google search intensity and weekly volatility. Thus, stocks tend to be more volatile when there is an increase in investor attention either in the current or preceding weeks (1-3 weeks).

Investor attention is significant at all lags when measured using the $AGSVI_{DA52,t}$ variable. It is noteworthy that the results are more robust when the Google search variable is lagged by three weeks, as regardless of how the Google variable is constructed ($AGSVI_{DA8}$; $AGSVI_{DA52}$ and $AGSVI_{BIJL52}$), the coefficients remain significant at the 1% and 5% levels. However, the impact on volatility diminishes as the AGSVI is lagged. For instance, one-unit increase in the $AGSVI_{DA52}$ variable at time t corresponds to 0.008 points of volatility increase in the same period. This effect diminishes to 0.0003 points of volatility when the variable is lagged by three weeks.

The results obtained for volatility in Model (1) align with those derived in the previous section concerning the return variable. An increase in investor attention, measured using the $AGSVI_{BIJL52}$ variable at $t-3$, has a significant impact on both volatility and returns in European markets.

Table 17. Regression results on Volatility with the AGSVI as independent variable.

	Volatility								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
AGSVI _{DA8,t}	0.006 (0.004)								
AGSVI _{DA52,t}		0.008** (0.004)							
AGSVI _{BIJL52,t}			0.001 (0.0003)						
AGSVI _{DA8,t-1}				0.003 (0.002)					
AGSVI _{DA52,t-1}					0.006*** (0.002)				
AGSVI _{BIJL52,t-1}						0.0003 (0.0002)			
AGSVI _{DA8,t-3}							0.003** (0.001)		
AGSVI _{DA52,t-3}								0.005*** (0.002)	
AGSVI _{BIJL52,t-3}									0.0003 (0.0001)
Observations	27,476	27,476	27,476	27,422	27,422	27,422	27,315	27,315	27,315
R ²	0.002	0.006	0.002	0.0004	0.003	0.0005	0.0004	0.002	0.001

Note: Models (1) to (3) assess the current relation of the AGSVI over volatility. Columns (4) to (6) display results of forecasting model of the AGSVI 1-week lagged. Columns (7) to (9) display results of forecasting model of the AGSVI 3-week lagged. Robust standard errors are clustered by ticker and time are reported in parentheses. The sample period covers weekly data from 2014 to 2024. The symbols ***, ** and * denote significance at the 1%, 5% and 10% levels, respectively.

Source: Own elaboration.

Next, Model (2) is conducted for the volatility specification. This autoregressive model extends Model (1) by including the weekly volatility of previous period in order to test whether the AGSVI can still capture the effect over volatility measure. Findings are presented in Table 18.

First, the exogenous variable of 1-week lagged volatility largely explains the fluctuations in the dependent variable, with the coefficient being approximately 0.512 (significant at the 1% level) across all model regressions. In other words, the previous week's volatility has a positive impact on the current period's volatility. This can be attributed to the nature of volatility, which tends to be persistent over time. Therefore, the lagged value is often a strong predictor of future volatility (Engle, 1982; Bollerslev, 1986; Adhikari and Hilliard, 2014)

When the AGSVI variable is 1-week lagged, the explanatory effect of the $AGSVI_{DA52}$ is significant at 5%, even when controlling for past volatility. The most robust effect is obtained when the regression is estimated with the AGSVI lagged by 3 weeks. The investor attention variable at $t-3$, regardless of how it is approached, has a significant relationship (at the 1%-5% level) on the current period's volatility. Therefore, the predictive power of Google searches persists and can explain some aspects of volatility behaviour that are not fully captured by past volatility. Furthermore, it is observed that regardless of the lag applied in the AGSVI, $AGSVI_{DA52}$ is the most outstanding indicator for capturing the effect on volatility.

Finally, Model (3) is tested using volatility variable. To examine whether the effect captured by Google search queries on volatility remain significant, control variables such as lagged return and trading volume have been included in the model. Table 18 presents the regression results of Model (3), with volatility as the dependent variable, and lagged volatility, trading volume, abnormal return and AGSVI as independent variables. Table 19 shows similar regressions, but considering excess return in place of abnormal return.

Both tables show that lagged volatility significantly explains current volatility at the 1% significance level. Additionally, an increase in the previous week's return negatively affects current volatility, with a stronger effect when measured using excess return. The coefficient with abnormal return is -0.028 (1% significant level) while in excess return -0.042 (1% significant level). This could be due to periods of higher returns are associated with a reduction in uncertainty and future volatility. In other words, a price increase may contribute to reducing the perception of risk.

Lastly, regarding the main variable of interest, the results indicate that an increase in search intensity at $t-3$ leads to an increase in current volatility (at the 1%-5% significance level). Therefore, past investor attention drives volatility, which could capture aspects not controlled by past return, volatility, or trading volume.

Table 18. Regression results on volatility with the AGSVI and lagged volatility as independent variable.

	Volatility								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
volatility _{t-1}	0.512*** (0.024)	0.510*** (0.024)	0.512*** (0.023)	0.513*** (0.023)	0.511*** (0.023)	0.513*** (0.023)	0.512*** (0.023)	0.511*** (0.023)	0.512*** (0.023)
AGSVI _{DA8,t}	0.0003 (0.002)								
AGSVI _{DA52,t}		0.002 (0.002)							
AGSVI _{BIJL52,t}			0.0001 (0.0002)						
AGSVI _{DA8,t-1}				-0.0003 (0.001)					
AGSVI _{DA52,t-1}					0.002** (0.001)				
AGSVI _{BIJL52,t-1}						-0.00000 (0.0001)			
AGSVI _{DA8,t-3}							0.002*** (0.001)		
AGSVI _{DA52,t-3}								0.003*** (0.001)	
AGSVI _{BIJL52,t-3}									0.0002** (0.0001)
Observations	27,419	27,419	27,419	27,419	27,419	27,419	27,312	27,312	27,312
R ²	0.263	0.263	0.263	0.263	0.263	0.263	0.263	0.263	0.263

Note: Models (1) to (3) assess the current relation of AGSVI and lagged volatility over Volatility. Columns (4) to (6) display results of forecasting model of AGSVI 1-week lagged. Columns (7) to (9) display results of forecasting model of AGSVI 3-week lagged. Robust standard errors are clustered by ticker and time are reported in parentheses. The sample period covers weekly data from 2014 to 2024. The symbols ***, ** and * denote significance at the 1%, 5% and 10% levels, respectively.

Source: Own elaboration.

The assessment of these research hypotheses is outlined below.

Hypothesis RH_{2a} proposed that the GSVI is positively associated with volatility. Overall, the results indicate that stocks tend to be more volatile when there is an increase in investor attention, either in the current or preceding weeks (1-3 weeks). In the models assessing the contemporaneous relationship between GSVI and volatility, statistical significance is observed only when the GSVI is measured using $AGSVI_{DA52}$ (at the 5% significance level), indicating limited robustness of the relationship. Furthermore, once lagged volatility is incorporated into Model (2), this significance disappears. Thus, while Model (1) suggests that a rise in search volume is positively linked to increased volatility, this relationship proves significant in only one of the three investor attention measurement approaches. Accordingly, hypothesis RH_{2a} is rejected, as the relationship is positive but lacks consistency across models.

Hypothesis RH_{2b} posits that the GSVI predicts stock volatility. To test this hypothesis, the effect of 1-3 weeks lagged $AGSVI$ is examined across three models, with stock volatility as the dependent variable. Both Model (1) and Model (2) indicate that an increase in investor demand is associated with a rise in volatility in the subsequent week. Similar to the contemporary relationship, this association only holds when investor attention is measured using $AGSVI_{DA52}$. However, when the Google variable is lagged by three weeks, the results reveal the predictive power of the GSVI on volatility, regardless the method used to calculate it. This result is consistent across all three models, remaining statistically significant at the 1%-5% level even after incorporating other financial variables in Model (3). Therefore, Hypothesis RH_{2b} is accepted.

Table 19. Regression results on volatility with lagged AR, volatility, ATV, DTV and the AGSVI.

	volatility											
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
AR _{t-1}	-0.028*** (0.004)	-0.028*** (0.004)	-0.028*** (0.004)	-0.028*** (0.004)	-0.028*** (0.004)	-0.028*** (0.004)	-0.028*** (0.004)	-0.028*** (0.004)	-0.028*** (0.004)	-0.028*** (0.004)	-0.028*** (0.004)	-0.028*** (0.004)
AGSVI _{DA8, t-1}	-0.0004 (0.001)			-0.0004 (0.001)								
AGSVI _{DA52, t-1}		0.001** (0.001)			0.002** (0.001)							
AGSVI _{BIJL52, t-1}			-0.00001 (0.0001)			-0.00001 (0.0001)						
AGSVI _{DA8, t-3}							0.002*** (0.001)			0.002*** (0.001)		
AGSVI _{DA52, t-3}								0.003*** (0.001)			0.003*** (0.001)	
AGSVI _{BIJL52, t-3}									0.0002** (0.0001)			0.0002** (0.0001)
Volatility _{t-1}	0.501*** (0.024)	0.500*** (0.024)	0.501*** (0.024)	0.504*** (0.027)	0.503*** (0.027)	0.503*** (0.027)	0.501*** (0.024)	0.500*** (0.024)	0.500*** (0.024)	0.503*** (0.027)	0.502*** (0.027)	0.503*** (0.027)
ATV _{t-1}	0.00004 (0.0001)	0.00004 (0.0001)	0.00004 (0.0001)				0.00004 (0.0001)	0.00004 (0.0001)	0.00004 (0.0001)			
DTV _{t-1}				-0.0002 (0.001)	-0.0003 (0.001)	-0.0003 (0.001)				-0.0003 (0.001)	-0.0002 (0.001)	-0.0002 (0.001)
Observations	27,284	27,284	27,284	27,419	27,419	27,419	27,187	27,187	27,187	27,312	27,312	27,312
R ²	0.267	0.267	0.267	0.267	0.267	0.267	0.267	0.267	0.267	0.266	0.267	0.266

Models (1) to (6) report the regression results of the 1-week lagged financial metrics and the AGSVI to forecast volatility. Models (7) to (12) report the regression results of the 1-week lagged financial metrics and the 3-weeks lagged AGSVI to forecast volatility. Robust standard errors are clustered by ticker and time are reported in parentheses. The sample period covers weekly data from 2014 to 2024. The symbols ***, ** and * denote significance at the 1%, 5% and 10% levels, respectively.

Source: Own elaboration.

Table 20. Regression results on volatility with lagged ER, volatility, ATV, DTV and the AGSVI.

	volatility											
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
ER _{t-1}	-0.042*** (0.004)	-0.042*** (0.004)	-0.042*** (0.004)	-0.042*** (0.004)	-0.042*** (0.004)	-0.042*** (0.004)	-0.042*** (0.004)	-0.042*** (0.004)	-0.042*** (0.004)	-0.042*** (0.004)	-0.042*** (0.004)	-0.042*** (0.004)
AGSVI _{DA8, t-1}	-0.0003 (0.001)			-0.0003 (0.001)								
AGSVI _{DA52, t-1}		0.002** (0.001)			0.002** (0.001)							
AGSVI _{BIJL52, t-1}			-0.00000 (0.0001)			-0.00000 (0.0001)						
AGSVI _{DA8, t-3}							0.002** (0.001)			0.002*** (0.001)		
AGSVI _{DA52, t-3}								0.003*** (0.001)			0.003*** (0.001)	
AGSVI _{BIJL52, t-3}									0.0002** (0.0001)			0.0002** (0.0001)
Volatility _{t-1}	0.496*** (0.024)	0.495*** (0.024)	0.496*** (0.024)	0.500*** (0.027)	0.499*** (0.027)	0.500*** (0.027)	0.496*** (0.024)	0.495*** (0.024)	0.496*** (0.024)	0.499*** (0.027)	0.498*** (0.027)	0.499*** (0.027)
ATV _{t-1}	0.00004 (0.0001)	0.00004 (0.0001)	0.00004 (0.0001)				0.00004 (0.0001)	0.00004 (0.0001)	0.00004 (0.0001)			
DTV _{t-1}				-0.0005 (0.001)	-0.001 (0.001)	-0.0005 (0.001)				-0.0005 (0.001)	-0.0004 (0.001)	-0.0005 (0.001)
Observations	27,284	27,284	27,284	27,419	27,419	27,419	27,187	27,187	27,187	27,312	27,312	27,312
R ²	0.269	0.269	0.269	0.269	0.270	0.269	0.269	0.270	0.269	0.269	0.269	0.269

Models (1) to (6) report the regression results of the 1-week lagged financial metrics and the AGSVI to forecast volatility. Models (7) to (12) report the regression results of the 1-week lagged financial metrics and the 3-weeks lagged AGSVI to forecast volatility. Robust standard errors are clustered by ticker and time are reported in parentheses. The sample period covers weekly data from 2014 to 2024. The symbols ***, ** and * denote significance at the 1%, 5% and 10% levels, respectively.

Source: Own elaboration.

6.2.3. GSVI and Trading Volume

The last result section presents the outcomes when testing the relation between investor attention variable and trading volume. The hypothesis RH_{3a} posed a positive relation between AGSVI and trading volume. To investigate whether investor attention drives trading volume, both abnormal and detrended trading volume are regressed against the current AGSVI. The hypotheses RH_{3b} puts forward the idea that Google searches predict trading activity. This approach allows to determine whether changes in investor focus, as indicated by search activity, have a delayed impact on trading volume.

As in previous sections, Table 21 presents a summary of the coefficients of AGSVI, the central variable of this research. The values corresponding to other variables included in the multiple regression model are omitted from this table, as their effects on the trading volume will be thoroughly examined subsequently in the section.

Table 21. Summary results on trading volume.

		ATV_t			DTV_t		
		t0	t-1	t-3	t0	t-1	t-3
<i>Model (1)</i>	AGSVI _{DA8,t}	0.255**	0.103	0.157**	0.135***	0.061***	0.033*
	AGSVI _{DA52,t}	0.332***	0.196**	0.179**	0.091***	0.028**	-0.002
	AGSVI _{BIJL52,t}	0.023	0.015*	0.016	0.007*	0.002	0.0001
<i>Model (2)</i>	AGSVI _{DA8,t}	0.234**	0.093	0.155**	0.046***	0.006	0.017
	AGSVI _{DA52,t}	0.311***	0.184**	0.174**	0.024**	-0.010	-0.005
	AGSVI _{BIJL52,t}	0.021	0.014*	0.015	0.002	-0.0005	-0.001
<i>Model (3)</i>	AR _t	AGSVI _{DA8,t}	0.019	0.133**		0.006	0.017
		AGSVI _{DA52,t}	0.077	0.110*		-0.010	-0.005
		AGSVI _{BIJL52,t}	0.007	0.012		-0.0005	-0.001

ER_t	$AGSVI_{DA8,t}$	0.019	0.131**	0.006	0.017
	$AGSVI_{DA52,t}$	0.079	0.111*	-0.009	-0.005
	$AGSVI_{BIJL52,t}$	0.007	0.012	-0.0005	-0.001

Note: Models (1) regresses the AGSVI against trading volume. Model (2) builds upon Model (1) by incorporating the lagged trading volume variable. Model (3) extends Model (2) by including lagged volatility and return. The sample period covers weekly data from 2014 to 2024. The symbols ***, ** and * denote significance at the 1%, 5% and 10% levels, respectively.

Source: Own elaboration.

The results presented in Tables 22 and 23 show the outcomes of the panel data regressions for Model (1), where abnormal trading volume and detrended trading volume are the dependent variables, and AGSVI (both in its current form and lagged by 1-3 weeks) is the independent variable. Overall, a positive relationship between investor attention and trading activity is consistently observed. Investors may be reacting to information or market signals, increasing investor attention, both in the current period and in the past, leading to higher trading volumes.

Regarding the forecasting approach, the findings indicate that an increase in investor attention predicts a rise in trading activity one and three weeks later. These results are similar when the trading activity is measured using ATV rather than DTV. It is worth noting that the coefficients are statistically significant at the 5% level when search activity is measured using $AGSVI_{DA52}$, both with 1-week and 3-week lags. This suggests that the other two AGSVI approaches are comparatively less robust. Furthermore, the predictive power persists across lags, although it tends to reduce progressively as the lag length increases.

Finally, when analysing the effect of investor attention on detrended trading volume, the results are only significant when using $AGSVI_{DA8}$ as the measure of investor attention. This suggests that the influence of past investor attention on trading volume weakens over time, and only certain measures of attention continue to have a significant impact.

Table 22. Regression results on abnormal trading volume with AGSVI as independent variable.

	ATV								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
AGSVI _{DA8,t}	0.255** (0.113)								
AGSVI _{DA52,t}		0.332*** (0.110)							
AGSVI _{BIJL52,t}			0.023 (0.014)						
AGSVI _{DA8,t-1}				0.103 (0.074)					
AGSVI _{DA52,t-1}					0.196** (0.081)				
AGSVI _{BIJL52,t-1}						0.015* (0.009)			
AGSVI _{DA8,t-3}							0.157** (0.066)		
AGSVI _{DA52,t-3}								0.179** (0.075)	
AGSVI _{BIJL52,t-3}									0.016 (0.010)
Observations	27,346	27,346	27,346	27,297	27,297	27,297	27,200	27,200	27,200
R ²	0.0001	0.0002	0.0001	0.00002	0.0001	0.00003	0.00004	0.0001	0.00004

Note: Models (1) to (3) assess the current relation of AGSVI over ATV. Columns (4) to (6) display results of forecasting model of AGSVI 1-week lagged. Columns (7) to (9) display results of forecasting model of AGSVI 3-week lagged. Robust standard errors are clustered by ticker and time are reported in parentheses. The sample period covers weekly data from 2014 to 2024. The symbols ***, ** and * denote significance at the 1%, 5% and 10% levels, respectively.

Source: Own elaboration.

Table 23. Regression results on detrended trading volume with AGSVI as independent variable.

	DTV								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
AGSVI _{DA8,t}	0.135** (0.028)								
AGSVI _{DA52,t}		0.091*** (0.017)							
AGSVI _{BIJL52,t}			0.007* (0.004)						
AGSVI _{DA8,t-1}				0.061*** (0.020)					
AGSVI _{DA52,t-1}					0.028** (0.013)				
AGSVI _{BIJL52,t-1}						0.002 (0.002)			
AGSVI _{DA8,t-3}							0.033* (0.018)		
AGSVI _{DA52,t-3}								-0.002 (0.014)	
AGSVI _{BIJL52,t-3}									0.0001 (0.001)
Observations	27,481	27,481	27,481	27,427	27,427	27,427	27,320	27,320	27,320
R ²	0.005	0.003	0.001	0.001	0.0003	0.0002	0.0003	0.00000	0.00000

Note: Models (1) to (3) assess the current relation of AGSVI over DTV. Columns (4) to (6) display results of forecasting model of AGSVI 1-week lagged. Columns (7) to (9) display results of forecasting model of AGSVI 3-week lagged. Robust standard errors are clustered by ticker and time are reported in parentheses. The sample period covers weekly data from 2014 to 2024. The symbols ***, ** and * denote significance at the 1%, 5% and 10% levels, respectively.

Source: Own elaboration.

Next, Model (2) adds the lagged trading volume variable to the regressions. Tables 24 and 25 present results with abnormal and detrended trading volume as dependent variable, respectively.

The nowcasting approach shows that an increase in Google search volume anticipates the current values of both abnormal and detrended trading volume. However, when the investor attention variable is added, the results differ

significantly depending on whether trading activity is measured using ATV or DTV.

Firstly, it is found a strong and significant coefficient when testing the predictive capacity of GSVI over abnormal trading volume, regardless of whether the lag is one or three weeks. An increase in search activity at time $t-1$ predicts a rise in trading activity when AGSVI is measured using $AGSVI_{DA52}$ and $AGSVI_{BIJL52}$, both statistically significant at the 5% level. These findings remain consistent with a 3-week lag, although in this case the significance holds for $AGSVI_{DA8}$ and $AGSVI_{DA52}$. Similar to the results in the volatility model and align with Model (1), $AGSVI_{DA52}$ emerges as the most robust of the three approaches for measuring investor attention.

The results are no longer statistically significant when trading activity is measured using DTV. This could be due to the fact that 1 -week lagged DTV explains current trading volume at a 1% significant level, thereby reducing the predictive power for the AGSVI variable.

Table 24. Regression results on abnormal trading volume with AGSVI ATV_{t-1} as independent variable.

	ATV								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
ATV_{t-1}	0.041 (0.029)	0.041 (0.029)	0.041 (0.029)	0.041 (0.029)	0.041 (0.029)	0.041 (0.029)	0.041 (0.029)	0.041 (0.029)	0.041 (0.029)
$AGSVI_{DA8,t}$	0.234** (0.096)								
$AGSVI_{DA52,t}$		0.331*** (0.100)							
$AGSVI_{BIJL52,t}$			0.021 (0.013)						
$AGSVI_{DA8,t-1}$				0.093 (0.065)					
$AGSVI_{DA52,t-1}$					0.184** (0.076)				
$AGSVI_{BIJL52,t-1}$						0.014* (0.008)			
$AGSVI_{DA8,t-3}$							0.155** (0.065)		

Note: Models (1) to (3) assess the current relation of AGSVI and lagged ATV over ATV. Columns (4) to (6) display results of forecasting model of AGSVI 1-week lagged. Columns (7) to (9) display results of forecasting model of AGSVI 3-week lagged. Robust standard errors are clustered by ticker and time are reported in parentheses. The sample period covers weekly data from 2014 to 2024. The symbols ***, ** and * denote significance at the 1%, 5% and 10% levels, respectively.

Table 25. Regression results on DTV volume with AGSVI and lagged DTV as independent variable.

[illegible]

Observations	27,427	27,427	27,427	27,427	27,427	27,427	27,320	27,320	27,320
R ²	0.169	0.168	0.168	0.168	0.168	0.168	0.168	0.168	0.168

Note: Models (1) to (3) assess the current relation of AGSVI and lagged DTV over DTV. Columns (4) to (6) display results of forecasting model of AGSVI 1-week lagged. Columns (7) to (9) display results of forecasting model of AGSVI 3-week lagged. Robust standard errors are clustered by ticker and time are reported in parentheses. The sample period covers weekly data from 2014 to 2024. The symbols ***, ** and * denote significance at the 1%, 5% and 10% levels, respectively.

Source: Own elaboration.

Finally, Model (3) is tested. Both financial metrics for trading volume are regressed against multiple lagged independent variables. Tables 26 and 27 display the results on the regressions for the multiple control variables combination.

Abnormal trading volume is driven by past volatility and return. An increase in past volatility is associated with an increase in trading volume at 1% significant level. In periods with risk or uncertainty (high volatility), investors tend to trade more, leading to higher trading activity. Contrary, the return is negatively related to trading volume. If returns are high and strong, investor would prefer to keep their positions decreasing its level of trading activity.

Regarding the investor attention variable, neither current nor 1-week lagged AGSVI captures abnormal trading volume movements. However, an increase on investor interest three weeks before impacts positively on abnormal trading volume, after controlling for volatility and return. This effect is only significant when proxying investor attention with AGSVI_{DA8} (5%) and AGSVI_{DA52} (1%).

When the trading activity is measured by the detrended variable, the results are different. The detrended trading volume is driven by past DTV and return, but not by volatility measure. Furthermore, it is not found any forecasting relation between investor attention and DTV.

After presenting in detail the results related to AGSVI and the trading volume variable, the evaluation of the previously stated research hypotheses is carried out.

Hypothesis RH_{3a} proposes the existence of a positive relationship between GSVI and trading volume in European markets. The results from Model (1) indicate

that an increase in search volume is positively associated with higher trading volume when investor attention is proxied by $AGSVI_{DA8}$ and $AGSVI_{DA52}$, with statistical significance at the 5% and 1% levels, respectively. Moreover, when the lagged abnormal trading volume is incorporated (Model 2), these results remain statistically significant. These findings are consistent when trading volume is measured using DTV. Investors are net buyers of attention-grabbing stocks and an increase in the attention demand (high GSVI) paid to the market will positively pressure the stocks liquidity, hence trading activity. Therefore, RH_{3a} is accepted.

Hypothesis RH_{3b} postulates that the GSVI can predict trading volume. For both ATV and DTV, an increase in $AGSVI_{t-1}$ is associated with a rise in trading activity one week later. However, the effect captured by AGSVI disappears once return and volatility are included as proposed in Model (3). In contrast, when AGSVI is lagged by 3 weeks, the results indicate that an increase in AGSVI predicts higher trading volume three weeks later, even when the dependent variable is controlled for return and volatility. This effect is only consistent when trading volume is measured using ATV. Therefore, the AGSVI appears to capture the effect on ATV more effectively than on DTV, being unusual deviations of trading activity better driven by significant increases in investor attention. All in all, the results show that an increase in information demand forecasts a subsequent rise in trading volume after 1 and 3 weeks. Consequently, RH_{3b} is accepted.

Table 26. Regression results on abnormal trading volume with lagged return, volatility, ATV and AGSVI.

	ATV											
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
ATV_{t-1}	0.031 (0.021)	0.031 (0.021)	0.031 (0.021)	0.031 (0.020)	0.031 (0.020)	0.031 (0.020)	0.031 (0.021)	0.031 (0.021)	0.031 (0.021)	0.031 (0.020)	0.031 (0.020)	0.031 (0.020)
$AGSVI_{DA8, t-1}$	0.019 (0.055)			0.019 (0.055)								
$AGSVI_{DA52, t-1}$		0.077 (0.050)			0.079 (0.051)							
$AGSVI_{BIJL52, t-1}$			0.007 (0.006)			0.007 (0.006)						
$AGSVI_{DA8, t-3}$							0.133** (0.058)			0.131** (0.058)		
$AGSVI_{DA52, t-3}$								0.110* (0.062)			0.111* (0.058)	
$AGSVI_{BIJL52, t-3}$									0.012 (0.009)			0.012 (0.009)
$Volatility_{t-1}$	12.734*** (1.079)	12.690*** (1.102)	12.718*** (1.082)	12.406*** (1.076)	12.360*** (1.100)	12.389*** (1.080)	12.735*** (1.082)	12.703*** (1.094)	12.725*** (1.080)	12.407*** (1.080)	12.375*** (1.093)	12.397*** (1.078)
AR_{t-1}	1.147*** (0.228)	1.146*** (0.228)	1.146*** (0.228)				1.144*** (0.228)	1.146*** (0.229)	1.146*** (0.228)			
ER_{t-1}				2.070*** (0.273)	2.070*** (0.273)	2.069*** (0.273)				2.066*** (0.273)	2.069*** (0.273)	2.069*** (0.273)
Observations	27,287	27,287	27,287	27,287	27,287	27,287	27,190	27,190	27,190	27,190	27,190	27,190
R ²	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006

Models (1) to (6) report the regression results of the 1-week lagged financial metrics and the AGSVI to forecast abnormal trading volume. Models (7) to (12) report the regression results of the 1-week lagged financial metrics and the 3-weeks lagged AGSVI to forecast abnormal trading volume. Robust standard errors are clustered by ticker and time are reported in parentheses. The sample period covers weekly data from 2014 to 2024. The symbols ***, ** and * denote significance at the 1%, 5% and 10% levels, respectively.

Source: Own elaboration.

Table 27. Regression results on DTV with lagged return, volatility, DTV and AGSVI

	DTV											
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
DTV_{t-1}	0.410*** (0.045)	0.410*** (0.045)	0.410*** (0.045)	0.407*** (0.045)	0.407*** (0.045)	0.407*** (0.045)	0.409*** (0.045)	0.409*** (0.045)	0.409*** (0.045)	0.406*** (0.045)	0.407*** (0.045)	0.407*** (0.045)
$AGSVI_{DA8,t-1}$	0.006 (0.012)			0.006 (0.012)								
$AGSVI_{DA52,t-1}$		-0.010 (0.009)			-0.009 (0.009)							
$AGSVI_{BIJL52,t-1}$			-0.0005 (0.001)			-0.0005 (0.001)						
$AGSVI_{DA8,t-3}$							0.017 (0.013)			0.017 (0.013)		
$AGSVI_{DA52,t-3}$								-0.005 (0.010)			-0.005 (0.010)	
$AGSVI_{BIJL52,t-3}$									-0.001 (0.001)			-0.001 (0.001)
$Volatility_{t-1}$	-0.105 (0.276)	-0.098 (0.276)	-0.103 (0.276)	-0.154 (0.274)	-0.148 (0.274)	-0.152 (0.274)	-0.097 (0.274)	-0.094 (0.272)	-0.095 (0.273)	-0.145 (0.273)	-0.142 (0.271)	-0.144 (0.272)
AR_{t-1}	-0.153*** (0.038)	-0.154*** (0.038)	-0.154*** (0.038)				-0.155*** (0.039)	-0.155*** (0.039)	-0.155*** (0.039)			
ER_{t-1}				-0.332*** (0.046)	-0.332*** (0.046)	-0.332*** (0.046)				-0.332*** (0.046)	-0.333*** (0.046)	-0.333*** (0.046)
Observations	27,422	27,422	27,422	27,422	27,422	27,422	27,315	27,315	27,315	27,315	27,315	27,315
R ²	0.169	0.169	0.169	0.170	0.170	0.170	0.168	0.168	0.168	0.170	0.170	0.170

Models (1) to (6) report the regression results of the 1-week lagged financial metrics and the AGSVI to forecast DTV. Models (7) to (12) report the regression results of the 1-week lagged financial metrics and the 3-weeks lagged AGSVI to forecast DTV. Robust standard errors are clustered by ticker and time are reported in parentheses. The sample period covers weekly data from 2014 to 2024. The symbols ***, ** and * denote significance at the 1%, 5% and 10% levels, respectively.

Source: Own elaboration.

VII – DISCUSSION

VII - DISCUSSION

This thesis is built on how investors perceive the information available in the market when making financial decisions and the way this process is connected to financial indicators that explain market behaviour.

The premise of investor attention theory is that investors have limited capacity to absorb and process information. As a result, individuals tend to focus only on a subset of available information while neglecting other relevant data (Da et al., 2011). In the realm of finance, this limitation forces investors to be selective in their investment choices, an idea that contrasts with traditional financial models.

Among these theories, the EMH (Fama, 1970) stands out, considering that stock markets are perfectly efficient and therefore incorporate all available information into asset prices. Assuming that investors are not able to reason and absorb all available information, due to limited cognitive constraints (Kahneman, 1973), this opens the door to questioning whether markets are truly efficient, or whether they may be significantly influenced by investors' bounded rationality, attention constraints and behavioural biases.

This thesis measures investor attention through the proxy variable GSVI, constructed from Google search activity. Since investors often seek information through online searches, this variable serves as an indicator of their investment intentions (Ramos et al., 2020; Swamy et al., 2019; Da et al., 2011). The main objective of this thesis is to examine the ability of the GSVI to anticipate and predict stock market movements in the European market. To this end, three pairs of hypotheses are proposed, introducing the GSVI variable into panel data regressions based on three dependent variables: return, volatility and trading volume. The anticipatory capacity of investor behaviour is evaluated using a nowcasting approach, which involves estimating current market movements based on contemporaneous investor attention (Cheraghali et al., 2023; Cho and Varian, 2012). This approach allows the assessment of how immediate changes in search behaviour are associated with ongoing fluctuations in stock market variables. In

turn, a forecasting approach is employed by lagging the GSVI variable, testing whether the effects of investor attention are not immediate but instead manifest in the market with a certain delay.

In line with the formulation of hypotheses and the presentation of results, this section discusses the findings within the broader theoretical framework and existing literature. Table 28 provides an overview of the study's hypotheses and their acceptance or rejection according to the results obtained. Prior to addressing on the main variables of the study, this section provides a discussion of the general findings.

Table 28. Hypothesis overview.

<i>RH_{1a}</i>	<i>GSVI is positively related with stock returns in European markets.</i>	<i>Rejected</i>
<i>RH_{1b}</i>	<i>GSVI predicts stock returns in European markets.</i>	<i>Partially Accepted</i>
<i>RH_{2a}</i>	<i>GSVI is positively related with stock volatility in European markets.</i>	<i>Rejected</i>
<i>RH_{2b}</i>	<i>GSVI predicts volatility in European markets.</i>	<i>Accepted</i>
<i>RH_{3a}</i>	<i>GSVI is positively related with stock trading volume in European markets.</i>	<i>Accepted</i>
<i>RH_{3b}</i>	<i>GSVI predicts trading volume in European markets.</i>	<i>Accepted</i>

Source: Own elaboration.

First, the analysis confirms that the GSVI is capable of anticipating trading volume (*RH_{3a}*), but not stock returns and volatility (*RH_{1a}* and *RH_{2a}*). This finding possesses that an increase on investor attention is closely linked to higher market participation activity. When investors search for financial information online, this typically reflects a growing interest regarding specific assets. This cognitive engagement often proceeds a decision to act, commonly through buy or sell transactions. In other words, trading volume serves as a natural channel through which investor attention materializes in the market. This association between the

GSVI and trading activity is consistent with previous studies in other markets, such as Desagre and D'Hondt (2019) and Da et al. (2011) in US or Adachi et al. (2017) in Asian markets.

Second, this thesis finds that the GSVI exhibits predictive power for stock returns, volatility and trading volume (RH_{1b} , RH_{2b} and RH_{3b}), indicating its potential to forecast future market behaviour. Moreover, the analysis indicates that the predictive power of the GSVI becomes more consistent and statistically significant as longer lags are applied. In other words, investor attention becomes more effective at triggering changes as the time window expands.

This finding suggests a temporal gap between information-seeking behaviour and actual market reactions, which may be attributed to two different ideas. On the one hand, it reinforces the ideal that information construction in markets does not happen instantly. Investors typically take time to react after obtaining new information, either because they need to process it or because they adopt gradual investment strategies. The results obtained from the GSVI reflect a process of lagged attention-driven behaviour, which may materialize in the market several weeks later.

On the other hand, based on the results, the market structure cause information to be incorporated into prices progressively, depending on liquidity and trading dynamic. Investor attention is a key input in price discovery process, but its impact depends on how efficiently the market is able to transform that attention into price movements. Less efficient markets, where frictions such as limited liquidity and heterogeneous investor behaviour prevail, attention tends to have a delayed but more persistent effect on market outcomes. This may signal growing investor engagement only materializes in trading activity and price adjustment over time.

In conclusion, it can be considered that this lagged attention reflects the information-processing dynamics of the market, capturing phases of assimilation and interpretation by investors. The following sections discusses the results based on the relationship of the GSVI with the variables of return, volatility and trading volume.

7.1. RETURN

To analyse the relationship between the GSVI and stock returns, this thesis proposed two hypotheses based on a nowcasting and forecasting approach. No clear trend has been observed over the years regarding the relationship between returns and the GSVI. Literature suggests that this relationship does not exhibit a consistent directional pattern, indicating that external factors or fluctuations in market conditions may influence the interaction between these variables. The following section presents the discussion of results based on the two hypotheses formulated regarding the effect of investor attention on stock returns.

The results of this study reject the first hypothesis RH_{1a} , as the relationship found between the GSVI and stock return is significant and negative only when the GSVI is estimated using $AGSVI_{DA52}$.

The literature remains inconclusive regarding the contemporaneous relationship between investor attention and stock returns. Research focusing specifically on European markets is limited. However, the findings of this thesis align with the existing studies conducted in such contexts. For instance, Ramos et al. (2020), who analysed the EUROSTOXX50 index but with a smaller sample of companies, did not observe any significant contemporaneous relationship between stock returns and Google search queries. Similarly, Kim et al. (2019), in their analysis of the Norwegian market, reported no clear evidence of a relationship, although they noted that this could be due to sample size limitations rather than the absence of a true effect.

Literature that has demonstrated a nowcasting approach using the GSVI suggests that stock returns are more closely linked to investor sentiment in response to specific events, rather than broader market dynamics. Notably, nowcasting captures short-term shifts in attention and sentiment, allowing for swift responses to emerging news, crisis, or awareness-drive topics. A sample of this can be seen in studies addressing socially or politically sensitive issues. For example, Ouadghiri et al. (2021) on sustainability, Piñeiro et al. (2020) on water-related companies, and Pereira et al. (2018) on investors responses to political events such as Donald Trump's presidency.

This thesis, however, analyses a 10-year period using weekly data, which reflects a more longitudinal behaviour trend in which short-term shocks are less pronounced. While studies focus on specific events, such as the COVID-19 crisis or awareness-related topics (Chundakkadan & Nedumparambil, 2022; Smales, 2021 and Ding et al., 2020), rely on daily or high frequency data to capture immediate investor reactions, their scope is often limited to exceptional circumstances. In contrast, the long-term perspective adopted in this thesis contributes to understanding structural patterns of investor behaviour in a more stable and mature markets such as those in Europe. These markets are characterized by lower volatility and stronger regulatory frameworks, making them well- suited for examining persistent relationships between investor attention and stock performance. Therefore, while short-term event driven studies capture sudden shifts in sentiment, this research provides evidence of a broader dynamics of investor attention over time.

Moving on to RH_{1b} , our results found that higher volumes of attention drive lower stock returns. However, although these results are consistent across models, it is significant in two of the three GSVI approaches. Hence, the RH_{1b} is partially accepted.

Literature broadly agrees that the GSVI has predictive power over stock returns. However, the evidence regarding the direction of this relationship remains inconclusive, as some studies find that the GSVI predicts lower returns, while others suggest it is related with future higher returns (Cheraghali et al., 2023; Swamy et al., 2019; Bijl et al., 2016 Da et al., 2011). The absence of consensus stems from the complex nature of stock returns, which makes it difficult to establish a clear relationship with investor attention.

One of the main results shows that 1-week lagged AGSVI variable is significantly associated with negative abnormal stock returns. Comparable outcomes are reported in Cheraghali et al. (2023), Yoshinaga and Rocco (2020) and Perlin et al. (2017), who found that high investor attention predicts lower stock returns one week after. Hence, not supporting the PPH, as an increase in investor interest is not followed by price pressure and positive future abnormal returns.

The predictive capacity of the GSVI is also tested with a three weeks lag. The analysis found that an increase in investor attention predicts lower stock returns

three weeks later. This result is consistent in the three models for both abnormal and excess return, when the GSVI is measured using the $AGSVI_{BIJL52}$ approach.

Interestingly, this finding is observed solely when returns are calculated using abnormal stock returns, as the significance of the GSVI coefficient vanishes when returns are calculated using excess return.

For further understand the nature of the relationship between investor attention and stock returns, this thesis uses two different return measures: abnormal and excess returns. While both serve as tools to isolate patterns in return behaviour, they differ in how they adjust for market effects and risk. Abnormal returns have been calculated relative to the benchmark Fama and French (1993) three factor model in order to capture firm-specific deviations from expected behaviour. In other words, this methodology reflects how much a stock's return differs from what would normally be anticipated due to events or news related directly to a firm. In contrast, excess returns are obtained by subtracting the risk-free rate from raw stock returns, adjusting from systematic risk, emphasising the relationship between investor behaviour and broader market dynamic, rather than company-specific movements.

The previous result suggests that the GSVI is more related to firm-specific factors rather than market-wide risk. The significant relationship between the GSVI and abnormal returns indicates that investor attention captured by Google searches may be driven by company-specific events. However, this effect disappears when using excess returns, which adjust for systematic risk. This indicates that the GSVI is primarily linked to abnormal and specific movements rather than broad market dynamics.

Another strand of the literature that adopts a forecasting approach using the GSVI concludes that higher levels of investor attention predict higher stock returns. Some examples can be shown in Adachi et al. (2017), Takeda and Wakao (2014), and Nguyen et al. (2020), who found that search intensity is positively associated with subsequent stock returns.

This thesis examines European markets, which are generally considered to be more traditional, stable, mature and regulated. These structural differences influence both investor behaviour and how attention translates into price movements. Thus, the previous studies must be interpreted within a different

market context, as they are based on structural conditions that diverge significantly from those analyzed in this thesis.

Firstly, the aforementioned research was conducted in Asian markets, where Google has lower market share and alternative search engines, such as Baidu, are more commonly used. This technological difference may directly affect the reliability of the GSVI as a proxy for investor attention in those settings.

Secondly, the nature of the markets themselves also differs considerably. For instance, Nguyen et al. (2020) focused on emerging economies, which are typically characterized by higher volatility and weaker regulatory frameworks. Similarly, Adachi et al. (2017) analysed a dataset composed of start-ups, representing a less mature and more speculative segment of the market. These factors can lead to different results of the effects of attention-driven behaviour on stock returns.

7.2. VOLATILITY

The second variable of study is volatility. Similar to the analysis conducted for stock returns a set of hypotheses is proposed to examine the effect of investor attention on market volatility within European markets, applying both a nowcasting and forecasting approach.

The results on RH_{2a} show that an increase in the GSVI explain current volatility only when the attention proxy is constructed using $AGSVI_{DA52}$. However, this effect vanishes once the lagged value of volatility is included as an explanatory variable. As a result, Hypothesis RH_{2a} is rejected. The literature on nowcasting volatility remains limited, as most studies tend to adopt a forecasting approach.

The results are consistent with previous findings such as Kim et al. (2019), who documented a contemporaneous relationship between Google search activity and market volatility in Norway. Their study suggested that increased online search behaviour reflects rising investor attention, increasing market volatility. However, aligning with the findings of Basistha et al. (2019) and supporting the results presented in this thesis, this relationship dissipates once lagged volatility is incorporated into the model.

This implies that lagged volatility already accounts for a substantial portion of the information that Google search activity would otherwise explain in isolation. This phenomenon can be understood in the context of financial time series dynamic, where the lagged value of volatility is widely recognized in the literature as one of the most robust predictors of future volatility (Engle, 1982; Bollerslev, 1986; Adhikari and Hilliard, 2014).

Regarding RH_{2b} , the results show that increased investor attention predicts higher stock market volatility, particularly when attention is lagged by three weeks. This relationship is statistically significant across all model specifications and GSVI construction methods. Therefore, RH_{2b} is accepted.

The predictive relationship identified in this thesis between Google search activity and stock volatility is in line with a broader consensus in the literature. Several studies across different markets report similar findings, confirming that investor attention serves as a consistent leading indicator of volatility. For example, Chen and Lo (2019) demonstrate this effect in the Taiwanese market, while Perlin et al. (2017) and Smales (2021) observe in a multi-country setting. Similarly, Hamid and Heiden (2015) find strong predictive power in the U.S, even showing that it can outperform traditional forecasting methods. These results are also in accordance with the findings of Vlastakis and Markellos (2012), who determined that an increase in information demand is significant and positively related to historical volatility, even after controlling for market returns.

Interestingly, this delayed effect of investor attention on volatility mirrors the results previously discussed in relation to stock returns. In both cases, the predictive power of the GSVI is strongest when the variable is lagged by three weeks, suggesting a consistent pattern in how markets process information. This reinforces the idea that investor and markets do not react to changes in public attention immediately, but rather with a certain delay.

The results on RH_{2b} clearly indicate that investor attention plays a crucial role in driving future volatility. When more attention is focused on a stock index, the amount of information uncovered and trading activity increases, leading prices to behave in a more random and less predictable manner. As more investors concentrate on a particular asset, price deviations from expected trends become more pronounced, causing higher future volatility. In this way, investor attention

can act as a mechanism that enhances price discovery and makes price movements more stochastic, thereby supporting market efficiency. This interpretation aligns with prior research by Tantaopas et al. (2016) and Vozlyublennaya (2014), who similarly argue that the positive relationship between attention and volatility contributes to market efficiency by promoting faster and more complete incorporation of information into prices.

However, it is also important to recognize that while increased attention can help information be reflected more quickly in prices, it may simultaneously encourage speculative behaviour and introduce noise. This can lead to short-term deviations from fundamental values and heightened volatility. Therefore, the observed link between attention and volatility does not necessarily imply greater market efficiency but rather a stronger market reaction. The outcome depends on the quality of the information attracting that attention.

7.3. TRADING VOLUME

Finally, the discussion extends to the relationship between investor attention and trading volume, following the same structure as in the analysis of stock returns and volatility.

Firstly, RH_{3a} was tested, confirming that a rise in investor attention is positively associated with an increase in trading volume in European markets. These results are in line with previous evidence from various markets, which consistently points to a positive link between investor attention and trading activity. Similar patterns have been observed by Kim et al. (2019) in Norway, Adachi et al. (2017) and Takeda and Wakao (2014) in Asian markets, and Bank et al. (2011) in Germany. Moreover, Desagre and D'Hondt (2019) confirm that this relationship holds even after accounting for sociodemographic and subjective investor characteristics, reinforcing the idea that increased search interest tends to translate into higher trading volume, regardless of regional or behavioural differences.

It is noteworthy to point out that, among the three hypotheses testing a nowcasting approach (stock returns, volatility, and trading volume) only the

relationship between investor attention and trading volume (RH_{3a}) is accepted based on the results. This highlights the fact that trading activity responds more immediately and directly to shifts in investor attention, as it reflects the actual decision to act on new information, regardless of whether prices or volatility have already adjusted.

When investors increase their demand for information, it clearly signals an intention to actively engage in trading related to the stock, resulting in higher trading volume. This occurs because heightened attention reflects both buying and selling interests. For example, Barber and Odean (2008) argue that increased investor attention is mainly linked to buying behaviour, as investors tend to seek new opportunities within stocks they already know. However, it is also evident that greater interest can drive selling activity, which similarly contributes to increased trading volume. Supporting this broader view, Desagre and D'Hondt (2019) challenge the PPH by showing that the relation between attention and volume is not stronger for purchases than for sales. Therefore, regardless of transaction type, increased investor attention is consistently and significantly associated with higher trading volume.

Regarding RH_{3b} , the results indicate that the GSVI with a 1–3-week lag effectively predicts trading volume. Specifically, the models using ATV yields a greater number of significant GSVI coefficients compared to DTV.

The fact that results differ depending on whether trading activity is measured using ATV or DTV suggests that unusual spikes in trading are more strongly influenced by sharp increases in investor attention. ATV captures deviations from a stock's typical trading behaviour relative to a historical benchmark, highlighting abnormal surges in activity. In contrast, DTV filters out long-term trends to isolate short-term fluctuations. The stronger and more consistent results found with ATV indicate that investor attention tends to drive trading behaviour particularly when it departs significantly from normal patterns, reinforcing the idea that attention is linked to extraordinary rather than normal market movements. In other words, when trading volume rises well above what is normally expected, it is likely being driven by a surge in investor attention. This suggests that people tend to search for information when something unusual is happening in the market, which then triggers more trading. Regular day-to-day trading, captured by DTV, may not be

as sensitive to changes in attention, but abnormal trading spikes seem closely tied to it.

There is a broad consensus in the literature regarding the predictive capacity of GSVI for abnormal trading activity. Sifat and Taker (2020) examined the predictive capacity of the GSVI in ASEAN markets, concluding that the GSVI has predictive power for high abnormal trading. Previously, Joseph et al. (2011) investigated the ability of search traffic to forecast ATV through a diversified portfolio model based on search intensity. The authors concluded that portfolios with higher search volume exhibited an average ATV 158% greater than portfolios with lower search levels. The predictive power of GSVI over trading volume is also supported by other studies, including Ramos et al. (2019), Kim et al. (2019), and Pinto et al. (2019), who find similar evidence across different market contexts. These findings reinforce the robustness of the forecasting relationship observed in this thesis.

Among the body of literature reviewed, only one study presents findings that diverge from the predominant consensus. Perlin et al. (2017) found that an increase in the GSVI related to financial terms such as "risk" and "debt" was associated with a decrease in trading volume. The authors suggested that this rise in the GSVI generated negative sentiment among investors, which in turn led to a reduction in trading activity for the stocks analysed. However, it is important to note that the GSVI in this study was constructed using stock tickers as the primary keywords. Given this methodological distinction, it can be inferred that the results in this study might differ, as there is an absence of inherent sentiment embedded in the context of the searched terms, unlike the emotionally charged financial keywords employed in the research by Perlin et al. (2017).

VI – CONCLUSIONS, LIMITATIONS AND FUTURE RESEARCH

VIII - CONCLUSIONS, LIMITATIONS AND FUTURE RESEARCH

This thesis measures investor attention using the GSVI, capturing a psychological concept that is otherwise difficult to quantify. Building on the theory of limited investor attention (Kahneman, 1973), which posits that investors selectively process information due to cognitive constraints, this work recognizes how such selective attention affects market behaviour and asset prices. Digital search data like the GSVI offers a novel, real-time proxy to empirically track these attention patterns.

The main contribution of this thesis lies in validating the GSVI as a reliable proxy for investor attention, particularly within stable markets. These markets, characterized by lower volatility, robust institutional frameworks, and stronger regulatory oversight, offer a suitable environment for examining more persistent and delayed relationships between attention and stock market outcomes. Within this context, the thesis focuses specifically on the European market. By applying nowcasting and forecasting techniques, the thesis demonstrates the predictive value of the GSVI-based indicator. The findings deepen theoretical understanding and suggest practical implications for using online search behaviour as a tool to anticipate market dynamics.

This chapter presents the main conclusions and contributions to the existing literature while synthesizing its theoretical and practical implications. Furthermore, this section discusses the limitations of the study and proposes future lines of research.

8.1. CONCLUSIONS AND IMPLICATIONS.

Firstly, this thesis significantly expands the conceptual framework surrounding investor attention by validating the GSVI as a credible and effective digital proxy for investor attention. By demonstrating that non-traditional data sources such as Google search queries reliably capture shifts in investor focus, this work supports their integration into behavioural finance models. It also highlights the critical role of varying levels of information accessibility and market frictions

in shaping investor behaviour and market outcomes, emphasizing that not all information is equally available or processed uniformly by market participants. While traditional finance theory assumes that all available information is instantly reflected in prices (Fama, 1970), empirical evidence shows that this is often not the case.

A key anomaly addressed in this thesis is rooted in the concept of limited attention (Shen and Wang, 2023; Chen et al., 2020; Kahneman 1973). Even when all investors have theoretical access to the same information, their cognitive capacity to process it is finite. As a result, they tend to concentrate only on the subset of information that is most salient or attention-grabbing. In this context, attention constraints function as a form of market friction, distorting information flow and introducing temporary mispricing. By using the GSVI as a tool to track this attention bias in real time, the thesis contributes to a better understanding of how behavioural patterns translate into observable market anomalies.

From a practical standpoint, the validation of the GSVI offers a novel and timely digital tool to monitor real-time investor attention. For instance, sudden spikes in the GSVI related to specific stocks or market events can serve as early warning signals of increased retail investor interest or emerging market trends. This capability allows for its incorporation into a broad range of forecasting models and sentiment-driven investment strategies, providing early signals of market engagement and shifts in investor interest. Financial analysts and asset managers can utilize the GSVI to detect changes in retail investor behaviour, such as increased searches ahead of earnings announcements or geopolitical events, which can inform risk assessment and support more informed asset allocation decisions. As a result, the GSVI stands out as a useful tool for portfolio managers navigating complex market dynamics driven by investor psychology.

Second, this thesis validates the relevance of the GSVI as a behavioural finance indicator tailored to the specific characteristics of European markets, enabling a more context-sensitive interpretation of investor attention. Despite growing academic interest in the use of the GSVI in markets such as the U.S. and Asia (Ouadghiri et al., 2021; Sifat and Thaker, 2020; Piñeiro et al., 2020; Nguyen et al., 2020), research applying this methodology to European stock markets remains scarce. This gap is particularly important given Europe's unique financial

environment, which includes a diverse investor base, varying regulatory frameworks, and specific market frictions that may influence how information is processed and acted upon by market participants.

This thesis highlights the need to rethink how investor attention is modelled in financial markets with structural characteristics like those found in Europe. Theoretically, the findings support the idea that behavioural responses are shaped by local market conditions. In particular, the slower information diffusion and lower volatility typically seen in European markets create a different environment for attention-driven behaviour compared to more reactive contexts such as the U.S. or Asia. This challenges the implicit assumption in much of the existing literature that behavioural finance mechanisms operate uniformly across markets. Instead, the results confirm that investor attention in Europe may build up more gradually and persist for longer periods, requiring models that account for these more subtle dynamics.

On the practical side, this thesis offers tools that are especially well-suited to the European financial context. For institutional investors and analysts operating in markets with strong regulatory oversight and relatively stable conditions (European Central Bank, 2025), tracking search volume can help identify early but consistent shifts in investor interest around key country-specific events, such as ECB announcements, national elections, or labour market data. In this sense, the GSVI provides a low-cost, real-time complement to traditional indicators, allowing for better timing in adjusting exposure and anticipating trends.

Moreover, in fragmented markets like those in Europe, where differences in language, regulation, and economic policy make cross-country analysis more complex, the use of search data offers a standardized way to gauge attention at the regional or national level. This is particularly valuable for investors managing diversified portfolios across European countries, as it helps detect where sentiment is intensifying and where it remains stable, thus informing more nuanced allocation decisions. All in all, this emphasizes the practical relevance of the GSVI as a tool for portfolio risk management, especially in markets defined by strong regulation and institutional stability.

Third, this thesis contributes by highlighting the limited effectiveness of the GSVI for nowcasting immediate market dynamics. The empirical analysis reveals

that, while sudden spikes in search activity reflect heightened investor attention and are often accompanied by increased trading volumes, they do not consistently correspond to simultaneous changes in asset prices or volatility. This difference can be explained by investor behaviour: trading volume reacts quickly to spikes in attention because active investors respond immediately, while price and volatility adjustments require broader information processing and consensus, causing a delay in their reaction. This suggests that investor attention, as captured by the GSVI, is not instantly absorbed into market behaviour, thereby limiting its utility for short-term forecasting.

From a theoretical standpoint, this supports behavioural finance theories that highlight cognitive limitations and the gradual diffusion of information. Even when investor attention rises sharply, market reactions may be delayed due to bounded rationality and the time it takes participants to process and act on new information. The observation that markets do not respond immediately to spikes in investor attention suggests that participants require time to interpret, evaluate, and act on new information. As a result, the impact of attention tends to emerge with a lag, rather than driving real-time price adjustments (Kim et al., 2019; Swamy et al., 2019; Perlin et al., 2017).

From a practical standpoint, this thesis highlights that while the GSVI has limited capacity to nowcast immediate market movements, it remains a valuable complementary indicator within broader analytical frameworks. Financial institutions and regulators could incorporate the GSVI into early warning systems to detect periods of market stress, thereby enhancing risk monitoring frameworks. Additionally, while the GSVI may not precisely nowcast price movements, it can help institutional traders anticipate changes in trading volume or spikes in market participation, which are crucial for optimizing order execution and liquidity management. Moreover, the GSVI is useful as an early indicator of shifts in investor engagement, contributing to more informed risk monitoring and decision-making. Furthermore, combining the GSVI with other measures like investor attention from social media can create a stronger overall sentiment index, making it easier to understand market feelings. Finally, since the GSVI likely reflects the behaviour of retail investors more strongly, it is valuable for identifying investor segments prone to herd behaviour or informational bubbles.

Fourth, the GSVI exhibits consistent predictive power that improves as the prediction horizon is extended. Specifically, forecasts with a three-week lag outperform those with a one-week lag, revealing that a longer time expand not only strengthens the signal but also offers significant practical advantages. This finding aligns closely with behavioural finance principles about how information spreads and how investors gradually process it over time (Thaler, 2016; Kahneman and Tversky, 1979). Investors do not absorb or act on all new information immediately. Instead, information diffuses slowly as investors interpret signals and adjust their positions at varying speeds.

The thesis demonstrates that lagged values of the GSVI are strongly and directly linked to changes in returns, volatility, and trading volume. It proves that online search behaviour reflects clear shifts in investor sentiment, which consistently take time to impact the market. This challenges the notion of strong market efficiency (Fama, 1970), which assumes prices instantly incorporate all relevant information. Instead, the evidence confirms that investor behaviour creates measurable delays between the availability of information and subsequent price adjustments.

From a practical standpoint, the enhanced predictive power of the GSVI at longer horizons is valuable for portfolio managers, risk analysts, and traders. Having a wider window to anticipate changes in returns, volatility, and trading volume allows market participants to move beyond reactive decisions. These lagged signals can help design better trading models that take into account how investor attention develops over time and affects different market variables. By observing spikes in investor attention (GSVI), they could identify potential periods of high volatility or uncertainty, enabling them to adjust portfolios to mitigate risks associated with sudden market movements or shifts in sentiment. This approach could ultimately lead to more informed, proactive strategies, improving the ability to capture returns and manage risk effectively. Moreover, longer prediction horizons help filter out short-term noise, providing more robust and stable signals that improve the reliability of forecasts and trading models. Ultimately, this means that GSVI transforms raw data on investor attention into actionable market intelligence with valuable lead time, empowering better decision-making and more effective navigation of market uncertainty.

Fifth, the construction of the GSVI plays a key role in its predictive performance. Specifically, the $AGSVI_{DA52}$ estimator shows stronger predictive ability for volatility and trading volume, whereas the $AGSVI_{BIJL52}$ estimator better explains variations in stock returns. This finding underscores the importance of carefully selecting the appropriate attention measure that aligns with the temporal dynamics and nature of the financial variable under analysis.

Choosing the right estimator is crucial because investor attention does not impact all market variables uniformly or instantaneously. Instead, its effects vary over time and depend on the specific characteristic being predicted. By recognizing these differences, researchers and practitioners can refine their forecasting models, leading to improved accuracy in predicting market behaviour.

From a practical perspective, this finding supports the integration of digital attention metrics into risk management systems and sentiment-driven trading strategies. For example, portfolio managers might rely on one estimator to anticipate upcoming volatility spikes while using another to forecast shifts in trading volume or price returns. This tailored approach enhances the development of more effective and responsive trading signals that better capture the complex and evolving nature of investor sentiment.

All in all, this thesis makes significant advances in understanding and utilizing investor attention through the GSVI. It clearly establishes that investor attention does not translate into immediate or uniform market reactions, supporting behavioural finance models that emphasize cognitive limitations and the gradual diffusion of information. This perspective is particularly important for European markets, where investor behaviour and information flow differ structurally from other regions.

The thesis confirms that the GSVI is a reliable, real-time digital indicator that captures early and measurable shifts in investor sentiment, shifts that become stronger and more useful when analysed over longer time horizons. This longer predictive window gives portfolio managers, risk analysts, and traders valuable advance notice of changes in returns, volatility, and trading volume, allowing them to act in a more deliberate and informed manner. Additionally, the identification of specific GSVI estimators tailored to different market variables improves the accuracy of forecasting models and trading strategies.

By combining novel digital data with behavioural finance concepts and empirical market evidence, this work provides market participants with a powerful tool to better understand and anticipate investor-driven market movements, ultimately supporting more effective risk management and improved investment decisions.

8.2. LIMITATION AND FUTURE RESEARCH

This study has contributed to understanding the relationship between investor attention and market behaviour. However, there are some areas of improvement and limitations that could be addressed in the future in order to extend and deepen the understanding of the topic. Acknowledging these limitations helps in guiding future research directions that could extend and deepen the understanding of the topic.

The use of regression-based approaches for econometric estimation is widely prevalent in the literature that examines the relationship between the GSVI and stock markets. For instance, the pooled linear panel model has been employed to study the impact of climate-related natural disasters on US stock indices (El Quadghiri, 2021). The VAR model has been used by researchers such as Padungsaksawasdi et al. (2019), Tantaopas et al. (2016), Perlin et al. (2017), and Dimpfl and Jank (2016). Additionally, panel data fixed effects models have been utilized by Cheraghali et al. (2022), Da et al. (2011), Smales et al. (2021), Bijl et al. (2016), and Kim et al. (2019) to further investigate the dynamics between GSVI and stock market behaviour. In this study, the fixed effects model has allowed for controlling unobserved heterogeneity between firms, ensuring consistent estimates of the relationship between the GSVI and returns, volatility, and trading volume. However, future studies could consider a VAR-based approach, which would allow for analysing how these variables interact dynamically over time. While the panel approach captures structural differences between firms, a VAR model could provide additional information on the short- and long-term effects between search volumes, returns, volatility, and trading volume.

A further limitation lies in the relatively recent availability of *Google Trends* as a data source. Given that the platform has only provided consistent and accessible data from 2004 onward, the time span available for analysis remains

limited when compared to more traditional financial or economic datasets. This constraint can reduce the size of the sample used in empirical models, particularly in studies involving long-term market trends or requiring extensive historical data for training and validation. While this does not undermine the value of the results obtained, it does suggest that conclusions drawn from Google Trends-based analyses should be interpreted within the context of the available timeframe. As the tool continues to evolve and accumulate more data over time, future studies will likely benefit from broader and more robust datasets.

It is also important to note that the predictive value of the GSVI is likely to improve as more historical search data becomes available and as analytical methods get better. With longer datasets and more refined models, future studies and practitioners may be able to extract even more accurate signals from online search behaviour. This means that digital attention metrics could become increasingly valuable tools in financial forecasting as the field continues to evolve.

Another potential limitation when employing *Google Trends* data lies in the fact that Google may update or modify its data processing methodology over time. These changes, such as adjustments in normalization, rounding, or filtering techniques, can influence how search volume is reported. Nonetheless, these changes are generally limited in scope and do not significantly compromise the overall reliability or interpretability of the data. While such updates are typically aimed at improving data quality or ensuring user privacy, they may inadvertently introduce inconsistencies when comparing data across different time periods. This could affect the historical continuity of the dataset, posing challenges for longitudinal analyses or models that rely on stable input variables. Consequently, any study using the GSVI should take into account the possibility that methodological changes in *Google Trends* could influence the robustness or replicability of its results.

Additionally, when using Google search data, it is often difficult to disentangle whether investor interest is driven by positive or negative sentiment, which may limit the interpretability of the results. While online searches are a valuable source of information regarding investor attention, they do not allow for a clear distinction between whether investors are reacting optimistically or pessimistically to market events. This lack of differentiation can complicate the

precise interpretation of investor behaviour. For instance, an increase in search volume could reflect interest driven by both positive news and concerns or uncertainties in the market, but it does not provide clear indications of whether the reactions are predominantly positive or negative. This limitation could lead to a biased interpretation of investor sentiment and decision-making processes.

Future research could focus on replicating similar experiments using the GSVI to predict behaviour in different financial markets. The results obtained in this study are clear and consistent for markets with similar characteristics, suggesting that subsequent investigations might find comparable patterns or even sector-specific outcomes. Additionally, it would be valuable to explore how the effectiveness of the GSVI varies across different regions or time periods, as well as to examine potential enhancements in integrating this variable with other economic indicators to improve market prediction accuracy.

Furthermore, future research could explore the use of lower-frequency data, such as daily data, to capture better investor intentions and market behaviour. While weekly frequency is widely used in financial studies due to its practicality and ability to smooth out short-term fluctuations, it may not fully capture the rapid absorption of price-related information or the immediate reactions to changes in investor sentiment. By utilizing daily data, researchers could gain a more detailed understanding of how quickly information is processed by the market, providing a clearer picture of short-term dynamics and investor behaviour. This approach could enhance the precision of findings and offer a more comprehensive view of the interactions between investor attention and market movements.

Looking ahead, the gradual nature of information diffusion underscores the importance of combining digital attention proxies like the GSVI with other financial indicators. This integration can help overcome the limitations of relying on any single measure and improve the robustness of predictive models. Moreover, as data availability and analytical methods evolve, the predictive utility of such proxies is likely to strengthen, opening new avenues for both research and practical application.

Finally, the growing reliance on online resources for financial decision-making may result in the faster dissemination of information, heightened market sensitivity, and a deeper understanding of the internet's impact on stock market

activities. Recognizing this dynamic could enable investors, analysts, and financial institutions to better adapt to the evolving digital landscape, enhancing their ability to predict market shifts and respond effectively to changes in investor sentiment driven by online activity.

IX – BIBLIOGRAPHICAL REFERENCES

IX - BIBLIOGRAPHICAL REFERENCES

- Acaravci, S. K., & Karaomer, Y. (2017). Fama-French five factor model: evidence from Turkey. *International Journal of Economics and Financial Issues*, 7(6), 130-137.
- Ackert, L. F., & Smith, B. F. (1993). Stock price volatility, ordinary dividends, and other cash flows to shareholders. *The Journal of Finance*, 48(4), 1147-1160.
- Ackert, L., & Deaves, R. (2009). Behavioural finance: Psychology, decision-making, and markets.
- Adachi Y, Masuda M, Takeda F (2017) Google search intensity and its relationship to the returns and liquidity of Japanese startup stocks. *Pac Basin Financ J* 46:243–257.
- Adhikari, B. K., & Hilliard, J. E. (2014). The VIX, VXO and realised volatility: A test of lagged and contemporaneous relationships. *International Journal of Financial Markets and Derivatives*, 3(3), 222-240.
- Ahundjanov, B. B., Akhundjanov, S. B., & Okhunjanov, B. B. (2020). Information search and financial markets under COVID-19. *Entropy*, 22(7), 791.
- Akarsu S, Süer Ö (2022) How investor attention affects stock returns? Some Int Evid *Borsa Istanb Rev* 22(3):616–626. <https://doi.org/10.1016/j.bir.2021.09.001>
- Alao, A. A. (2018). Issues in Information Asymmetries and Financial Markets: A Review of Literature. *Journal of Accounting and Financial Management* ISSN, 4(2), 2018.
- Alberg, D., Shalit, H., & Yosef, R. (2008). Estimating stock market volatility using asymmetric GARCH models. *Applied Financial Economics*, 18(15), 1201-1208.
- Amini, S., Delgado, M. S., Henderson, D. J., & Parmeter, C. F. (2012). Fixed vs random: The Hausman test four decades later. In *Essays in honor of Jerry Hausman* (Vol. 29, pp. 479-513). Emerald Group Publishing Limited.
- Andersen, J. V. (2010). Detecting anchoring in financial markets. *Journal of Behavioural finance*, 11(2), 129-133.

- Andersen, T. G. (1996). Return volatility and trading volume: An information flow interpretation of stochastic volatility. *The Journal of Finance*, 51(1), 169-204
- Anderson, A. M., & Dyl, E. A. (2005). Market structure and trading volume. *Journal of Financial Research*, 28(1), 115-131.
- Ané, T., & Ureche-Rangau, L. (2008). Does trading volume really explain stock returns volatility?. *Journal of International Financial Markets, Institutions and Money*, 18(3), 216-235.
- Aouadi A, Arouri M, Teulon F (2013) Investor attention and stock market activity: evidence from France. *Econ Model* 35:674–681.
- Aouadi, Amal; Arouri, Mohamed; Roubaud, David (2017). Information demand and stock market liquidity: International evidence. *Economic Modelling*, (), S0264999317316589–.
- Araújo RDA, Oliveira AL, Meira S (2015) A hybrid model for high-frequency stock market forecasting. *Expert Syst Appl* 42(8):4081–4096.
- Arditi E, Yechiam E, Zahavi G (2015) Association between stock market gains and losses and Google searches. *PLoS ONE* 10(10):e0141354.
- Artmann, S., Finter, P., & Kempf, A. (2012). Determinants of expected stock returns: large sample evidence from the German market. *Journal of Business Finance & Accounting*, 39(5-6), 758-784.
- Ayala, M.J., González-Gallego, N. & Arteaga-Sánchez, R. Google search volume index and investor attention in stock market: a systematic review. *Financ Innov* 10, 70 (2024).
- 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.
- Ayers, J. W., Althouse, B. M., Allem, J. P., Rosenquist, J. N., & Ford, D. E. (2013). Seasonality in seeking mental health information on Google. *American journal of preventive medicine*, 44(5), 520-525.
- Baker, M., & Wurgler, J. (2007). Investor sentiment in the stock market. *Journal of economic perspectives*, 21(2), 129-151.

- Bali, T. G., & Peng, L. (2006). Is there a risk–return trade-off? Evidence from high-frequency data. *Journal of Applied Econometrics*, 21(8), 1169-1198.
- Bank M, Larch M, Peter G (2011) Google Search volume and its influence on liquidity and returns of German stocks. *Financ Mark Portf Manag* 24(3):239–264. <https://doi.org/10.1007/s11408-011-0165-y>
- Bantis, E., Clements, M. P., & Urquhart, A. (2023). Forecasting GDP growth rates in the United States and Brazil using Google Trends. *International Journal of Forecasting*, 39(4), 1909-1924.
- Banz, R. W. (1981). The relationship between return and market value of common stocks. *Journal of financial economics*, 9(1), 3-18
- Barber BM, Odean T (2008) All that glitters: the effect of attention and news on the buying behaviour of individual and institutional investors. *Rev Financ Stud* 21(2):785–818.
- Basistha A, Kurov A, Wolfe M (2019) Volatility forecasting: the role of internet search activity and implied volatility. <https://doi.org/10.2139/ssrn.2812387>
- Batrinca, B., Hesse, C. W., & Treleven, P. C. (2018). Examining drivers of trading volume in European markets. *International Journal of Finance & Economics*, 23(2), 134-154.
- Bazán-Palomino W, Svogun D (2023) On the drivers of technical analysis profits in cryptocurrency markets: a Distributed Lag approach. *Int Rev Financ Anal*.
- Ben-Rephael, A., Da, Z., & Israelsen, R. D. (2017). It depends on where you search: Institutional investor attention and underreaction to news. *The Review of Financial Studies*, 30(9), 3009-3047.
- Bijl L, Kringhaug G, Molnár P, Sansvik E (2016) Google searches and stock returns. *Int Rev Financ Anal* 45:150–156.
- Bloom, R., Amber, K. T., Hu, S., & Kirsner, R. (2015). Google search trends and skin cancer: evaluating the US population's interest in skin cancer and its association with melanoma outcomes. *JAMA dermatology*, 151(8), 903-905.
- Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics*, 31(3), 307-327.

- Bonfiglioli, A., & Favero, C. A. (2005). Explaining co-movements between stock markets: The case of US and Germany. *Journal of International Money and Finance*, 24(8), 1299-1316.
- Bordalo, P., Gennaioli, N., & Shleifer, A. (2012). Salience theory of choice under risk. *The Quarterly journal of economics*, 127(3), 1243-1285.
- Bordalo, P., Gennaioli, N., & Shleifer, A. (2022). Salience. *Annual Review of Economics*, 14(1), 521-544.
- Bordalo, Pedro, Nicola Gennaioli, Rafael La Porta, and Andrei Shleifer. 2019. "Diagnostic Expectations and Stock Returns." *Journal of Finance* 74 (6): 2839-74.
- Borup, D., & Schütte, E. C. M. (2022). In search of a job: Forecasting employment growth using Google Trends. *Journal of Business & Economic Statistics*, 40(1), 186-200.
- Bose, S. (2007). Contribution of Indian index futures to price formation in the stock market. *Money & Finance*, 3(1).
- Brandt, M. W., & Kinlay, J. (2005). Estimating historical volatility. Available at SSRN 4384038.
- Broder A (2002) A taxonomy of web search. *SIGIR Fórum* 36(2):3-10.
- Brooks, C. (2019). *Introductory econometrics for finance*. Cambridge university press
- Bustos O, Pomares-Quimbaya A (2020) Stock market movements forecast: a systematic review. *Expert Syst Appl* 156(20):113464.
- Campbell, J. Y., Grossman, S. J., & Wang, J. (1993). Trading volume and serial correlation in stock returns. *The Quarterly Journal of Economics*, 108(4), 905-939.
- Campbell, S. D., & Sharpe, S. A. (2009). Anchoring bias in consensus forecasts and its effect on market prices. *Journal of Financial and Quantitative Analysis*, 44(2), 369-390.
- Cardoso-Leite, P., Green, C. S., & Bavelier, D. (2015). On the impact of new technologies on multitasking. *Developmental Review*, 35, 98-112.
- Carhart, M. (1997). On persistence in Mutual Fund Performance. *Journal of Financial Economics*, Vol. 52, No. 1, 57-82.

- CEIC Data. (2023). Frankfurt Stock Exchange domestic market cap. CEIC Data. <https://www.ceicdata.com>
- Chai, D., Faff, R., & Gharghori, P. (2010). New evidence on the relation between stock liquidity and measures of trading activity. *International Review of Financial Analysis*, 19(3), 181-192.
- Challet D, Bel Hadj Ayed A (2013) Predicting financial markets with Google trends and not so random keywords. In: Browser download this paper. <https://doi.org/10.48550/arXiv.1307.4643>
- Chaudhary, R., Bakhshi, P., & Gupta, H. (2020). Volatility in international stock markets: An empirical study during COVID-19. *Journal of Risk and Financial Management*, 13(9), 208.
- Chemmanur, T. J., & Yan, A. (2019). Advertising, attention, and stock returns. *Quarterly Journal of Finance*, 9(03), 1950009.
- Chen J, Tang G, Yao J, Zhou G (2022) Investor attention and stock returns. *J Financ Quant Anal* 57(2):455–484.
- Chen, H. Y., & Lo, T. C. (2019). Online search activities and investor attention on financial markets. *Asia Pacific Management Review*, 24(1), 21-26.
- Chen, H., Harrison, D. M., & Khoshnoud, M. (2020). Investors' limited attention: Evidence from REITs. *The Journal of Real Estate Finance and Economics*, 61(3), 408-442.
- Chen, S. S. (2012). Revisiting the empirical linkages between stock returns and trading volume. *Journal of Banking & Finance*, 36(6), 1781-1788.
- Chen, T. C., & Lai, M. Y. (2013). Are investors rational? Evidence on the impact of industrial framing reclassification on stock market reaction. *Journal of Behavioural Finance*, 14(1), 1-8.
- Cheraghali, H., Høydal, H., Lysebo, C., & Molnár, P. (2023). Consumer attention and company performance: Evidence from luxury companies. *Finance Research Letters*, 58, 104280.
- Chiang, T. C., Qiao, Z., & Wong, W. K. (2010). New evidence on the relation between return volatility and trading volume. *Journal of Forecasting*, 29(5), 502-515.

- Choi, J. J., Hauser, S., & Kopecky, K. J. (1999). Does the stock market predict real activity? Time series evidence from the G-7 countries. *Journal of Banking & Finance*, 23(12), 1771-1792.
- Chordia, T., Roll, R., & Subrahmanyam, A. (2011). Recent trends in trading activity and market quality. *Journal of Financial Economics*, 101(2), 243-263.
- Chronopoulos DK, Papadimitriou FI, Vlastakis N (2018) Information demand and stock return predictability. *J Int Money Financ* 80:59–74.
- Chuang, W. I., Liu, H. H., & Susmel, R. (2012). The bivariate GARCH approach to investigating the relation between stock returns, trading volume, and return volatility. *Global Finance Journal*, 23(1), 1-15.
- Chundakkadan, R., & Nedumparambil, E. (2022). In search of COVID-19 and stock market behaviour. *Global Finance Journal*, 54, 100639.
- Chung, S. L., Hung, C. H., & Yeh, C. Y. (2012). When does investor sentiment predict stock returns?. *Journal of Empirical Finance*, 19(2), 217-240.
- Cieslak, Anna. 2018. "Short-Rate Expectations and Unexpected Returns in Treasury Bonds." *Review of Financial Studies* 31 (9): 3265–3306.
- Clark, P. K. (1973). A subordinated stochastic process model with finite variance for speculative prices. *Econometrica: journal of the Econometric Society*, 135-155.
- Clarke, J., Chen, H., Du, D., & Hu, Y. J. (2020). Fake news, investor attention, and market reaction. *Information Systems Research*, 32(1), 35-52.
- Cohen, G., & Kudryavtsev, A. (2012). Investor rationality and financial decisions. *Jo*
- Coibion, O., & Gorodnichenko, Y. (2015). Is the Phillips curve alive and well after all? Inflation expectations and the missing disinflation. *American Economic Journal: Macroeconomics*, 7(1), 197-232.
- Coibion, O., Gorodnichenko, Y., & Kamdar, R. (2018). The formation of expectations, inflation, and the phillips curve. *Journal of Economic Literature*, 56(4), 1447-1491.
- Cready, W. M., & Gurun, U. G. (2010). Aggregate market reaction to earnings announcements. *Journal of Accounting Research*, 48(2), 289-334.

- D'Amuri, F., & Marcucci, J. (2017). The predictive power of Google searches in forecasting US unemployment. *International Journal of Forecasting*, 33(4), 801-816.
- Da Z, Engelberg J, Gao P (2011) In search of attention. *J Financ* 66(5):1461–1499.
- Darolles, S., Le Fol, G., & Mero, G. (2015). Measuring the liquidity part of volume. *Journal of Banking & Finance*, 50, 92-105.
- Dávila, E., & Parlato, C. (2021). Trading costs and informational efficiency. *The Journal of Finance*, 76(3), 1471-1539.
- De Bondt, W. F., & Thaler, R. (1985). Does the stock market overreact?. *The Journal of finance*, 40(3), 793-805.
- de Mendoza, F. P. (2008). *Dirección financiera*. Civitas.
- Debata, B., Dash, S. R., & Mahakud, J. (2018). Investor sentiment and emerging stock market liquidity. *Finance Research Letters*, 26, 15-31.
- Desagre, C., & D'Hondt, C. (2018). *Googlization and retail investment decisions* (No. UCL-Université Catholique de Louvain).
- Devenow, A. and Welch, I. (1996), "Rational herding in financial economics", *European Economic Review*, Vol. 40 Nos 3/5, pp. 603-615
- Dharani, M., Hassan, M. K., Abedin, M. Z., & Ismail, M. A. (2022). Does a search attention index explain portfolio returns in India?. *Borsa Istanbul Review*, 22(2), 226-239.
- Dhingra, B., Batra, S., Aggarwal, V., Yadav, M., & Kumar, P. (2024). Stock market volatility: a systematic review. *Journal of Modelling in Management*, 19(3), 925-952.
- Dimpfl T, Jank S (2016) Can internet search queries help to predict stock market volatility? *Eur Financ Manag* 22(2):171–192.
- Ding D, Guan C, Chan CM, Liu W (2020) Building stock market resilience through digital transformation: using Google trends to analyze the impact of COVID-19 pandemic. *Front Bus Res China* 14(1):1–21.
- Ding, Rong; Hou, Wenxuan (2015). Retail investor attention and stock liquidity. *Journal of International Financial Markets, Institutions and Money*, 37(), 12–26.

- Dirkx, P., & Peter, F. J. (2020). The Fama-French five-factor model plus momentum: Evidence for the German market. *Schmalenbach Business Review*, 72, 661-684.
- Doyne Farmer 5, J., Gillemot, L., Lillo, F., Mike, S., & Sen, A. (2004). What really causes large price changes?. *Quantitative finance*, 4(4), 383-397.
- Dzielinski M (2012) Measuring economic uncertainty and its impact on the stock market. *Finance Res Lett* 9(3):167–175. <https://doi.org/10.1016/j.frl.2011.10.003>
- Ekinci C, Bulut AE (2021) Google search and stock returns: a study on BIST 100 stocks. *Glob Financ J* 47:100518. <https://doi.org/10.1016/j.gfj.2020.100518>
- El Ouadghiri I, Guesmi K, Peillex J, Ziegler A (2021) Public attention to environmental issues and stock market returns. *Ecol Econ* 180:106836. <https://doi.org/10.1016/j.ecolecon.2020.106836>
- El Ouadghiri I, Peillex J (2018) Public attention to “Islamic terrorism” and stock market returns. *J Comp Econ* 46(4):936–946. <https://doi.org/10.1016/j.jce.2018.07.014>
- Elton, E. J., Gruber, M. J., & Busse, J. A. (2004). Are investors rational? Choices among index funds. *the Journal of Finance*, 59(1), 261-288.
- Elton, E. J., Gruber, M. J., Das, S., & Hlavka, M. (1993). Efficiency with costly information: A reinterpretation of evidence from managed portfolios. *The review of financial studies*, 6(1), 1-22.
- Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica: Journal of the econometric society*, 987-1007.
- Euronext. (n.d.). Euronext confirms its position as number one equity listing venue. Euronext. Retrieved July 23, 2024, from <https://www.euronext.com/en/about/media/euronext-press-releases/euronext-confirms-its-position-number-one-equity-listing-venue>
- European Central Bank. (2025, May). *Financial stability review – May 2025*. <https://www.ecb.europa.eu/press/financial-stability-publications/fsr/html/ecb.fsr202505~0cde5244f6.en.html>

- Evans, M. D., & Hnatkovska, V. V. (2014). International capital flows, returns and world financial integration. *Journal of International Economics*, 92(1), 14-33.
- Fama, E. F. (1970). Efficient capital markets. *Journal of finance*, 25(2), 383-417.
- Fama, E. F. (1970). Session topic: stock market price behaviour. *The Journal of Finance*, 25(2), 383-417.
- Fama, E. F. (1995). Random walks in stock market prices. *Financial analysts journal*, 51(1), 75-80.
- Fama, E. F., and K. R. French (1993), 'Common Risk Factors in the Returns on Stock and Bonds', *Journal of Financial Economics*, 33(1), 3-56.
- Fama, E. F., and K. R. French (2015), 'A Five-Factor Asset Pricing Model', *Journal of Financial Economics*, 116(1), 1-22.
- Fan, M. H., Chen, M. Y., & Liao, E. C. (2021). A deep learning approach for financial market prediction: Utilization of Google trends and keywords. *Granular Computing*, 6(1), 207-216.
- Ferrara, L., & Simoni, A. (2023). When are Google data useful to nowcast GDP? An approach via preselection and shrinkage. *Journal of Business & Economic Statistics*, 41(4), 1188-1202.
- Filippou IT, Nguyen M, Taylor MP (2023) Investor attention to news on financial integration and currency returns. In: Mark P. (Eds.) *Investor attention to news on financial integration and currency returns* (February 20, 2023). <https://doi.org/10.2139/ssrn.4364178>
- Fink, C., & Johann, T. (2014). May I have your attention, please: The market microstructure of investor attention. *Please: The Market Microstructure of Investor Attention* (September 17, 2014)
- Forsythe, R., Rietz, T. A., & Ross, T. W. (1999). Wishes, expectations and actions: a survey on price formation in election stock markets. *Journal of Economic Behaviour & Organization*, 39(1), 83-110.
- Ganesh, A., & Iyer, S. (2023). Impact of firm-initiated tweets on stock return and trading volume. *Journal of Behavioural Finance*, 24(2), 171-182.
- Gao, Y., Li, Y., & Wang, Y. (2023). The dynamic interaction between investor attention and green security market: an empirical study based on Baidu index. *China Finance Review International*, 13(1), 79-101.

- Garman, M. B., & Klass, M. J. (1980). On the estimation of security price volatilities from historical data. *Journal of business*, 67-78.
- Ghaddab, S., de Peretti, C., & Belkacem, L. (2024). Are stock markets efficient with respect to the Google Search Volume Index? A robustness check of the literature studies. *Research in International Business and Finance*, 102574.
- Ghysels, E., Santa-Clara, P., & Valkanov, R. (2005). There is a risk-return trade-off after all. *Journal of financial economics*, 76(3), 509-548.
- Gkillas, K., Vortelinos, D. I., & Saha, S. (2018). The properties of realized volatility and realized correlation: Evidence from the Indian stock market. *Physica A: Statistical Mechanics and its Applications*, 492, 343-359.
- Glaser, M., Langer, T., Reynders, J., & Weber, M. (2007). Framing effects in stock market forecasts: The difference between asking for prices and asking for returns. *Review of Finance*, 11(2), 325-357.
- Gödker, K., & Lukas, M. (2021). Attention to extreme returns. *Available at SSRN* 3080332.
- Goel, A., Tripathi, V., & Agarwal, M. (2021). Information asymmetry and stock returns. *Journal of Advances in Management Research*, 18(1), 85-112.
- Gómez-Martínez, R., Orden-Cruz, C., & Martínez-Navalón, J. G. (2022). Wikipedia pageviews as investors' attention indicator for Nasdaq. *Intelligent systems in accounting, finance and management*, 29(1), 41-49.
- González-Velasco C, González-Fernández M (2023) Has COVID-19 pandemic fear affected eurozone stock markets? Crises Uncertain Econ. https://doi.org/10.1007/978-981-19-3296-0_14
- Gonzálvez-Gallego, N., Pérez-Cárceles, M. C., & Nieto-Torrejón, L. (2024). Do search queries predict violence against women? A forecasting model based on Google Trends. *Journal of Forecasting*.
- Götz, T. B., & Knetsch, T. A. (2019). Google data in bridge equation models for German GDP. *International Journal of Forecasting*, 35(1), 45-66
- Gourieroux, C., Jasiak, J., & Le Fol, G. (1999). Intra-day market activity. *Journal of financial markets*, 2(3), 193-226.
- Greenwood, Robin, and Andrei Shleifer. 2014. "Expectations of Returns and Expected Returns." *Review of Financial Studies* 27 (3): 714–46.

- Griffin, J. M., Nardari, F., & Stulz, R. M. (2007). Do investors trade more when stocks have performed well? Evidence from 46 countries. *The Review of Financial Studies*, 20(3), 905-951.
- Gu, B., & Hitt, L. (2001). Transaction costs and market efficiency. *ICIS 2001 Proceedings*, 11.
- Hamid A, Heiden M (2015) Forecasting volatility with empirical similarity and Google Trends. *J Econ Behav Organ* 117:62–81. <https://doi.org/10.1016/j.jebo.2015.06.005>
- Hausman, J. A. (1978). Specification tests in econometrics. *Econometrica: Journal of the econometric society*, 1251-1271.
- Heiberger RH (2015) Collective attention and stock prices: evidence from Google Trends data on standard and poor's 100. *PLoS ONE* 10(8):e0135311. <https://doi.org/10.1371/journal.pone.0135311>
- Herawati, A., & Putra, A. S. (2018). The influence of fundamental analysis on stock prices: The case of food and beverage industries. *European Research Studies*, 21(3), 316-326.
- Herwartz, H. (2017). Stock return prediction under GARCH—An empirical assessment. *International Journal of Forecasting*, 33(3), 569-580.
- Heston, S. L., Korajczyk, R. A., & Sadka, R. (2010). Intraday patterns in the cross-section of stock returns. *The Journal of Finance*, 65(4), 1369-1407.
- Hirshleifer, D., Lim, S. S., & Teoh, S. H. (2009). Driven to distraction: Extraneous events and underreaction to earnings news. *The journal of finance*, 64(5), 2289-2325.
- Hjalmarsson, E. (2010). Predicting global stock returns. *Journal of Financial and Quantitative Analysis*, 45(1), 49-80.
- Hopkins, Z. H., & Secrest, A. M. (2019). Public health implications of Google searches for sunscreen, sunburn, skin cancer, and melanoma in the United States. *American Journal of Health Promotion*, 33(4), 611-615.
- Hou, K., Peng, L., & Xiong, W. (2009). A tale of two anomalies: The implications of investor attention for price and earnings momentum. *Available at SSRN* 976394.

- Hou, Kewei, Lin Peng, and Wei Xiong. "A Tale of Two Anomalies: The Implication of Investor Attention for Price and Earnings Momentum." (2006).
- Hu, J., Sui, Y., & Ma, F. (2021). The measurement method of investor sentiment and its relationship with stock market. *Computational Intelligence and Neuroscience*, 2021(1), 6672677.
- Huang MY, Rojas RR, Convery PD (2020) Forecasting stock market movements using Google Trend searches. *Empir Econ* 59(6):2821–2839. <https://doi.org/10.1007/s00181-019-01725-1>
- Huang, M. Y., Rojas, R. R., & Convery, P. D. (2020). Forecasting stock market movements using Google Trend searches. *Empirical Economics*, 59, 2821-2839.
- Huang, T. L. (2019). Is the Fama and French five-factor model robust in the Chinese stock market?. *Asia Pacific Management Review*, 24(3), 278-289.
- Iliev, P., Kalodimos, J., & Lowry, M. (2021). Investors' attention to corporate governance. *The Review of Financial Studies*, 34(12), 5581-5628.
- Izzeldin, M., Muradoğlu, Y. G., Pappas, V., & Sivaprasad, S. (2021). The impact of Covid-19 on G7 stock markets volatility: Evidence from a ST-HAR model. *International Review of Financial Analysis*, 74, 101671.
- Jacobsen, B., Mamun, A., & Visaltanachoti, N. (2005). Seasonal, size and value anomalies. *Available at SSRN* 784186.
- Jensen, M. C. (1978). Some anomalous evidence regarding market efficiency. *Journal of financial economics*, 6(2/3), 95-101.
- Jiang, G. J., & Tian, Y. S. (2005). The model-free implied volatility and its information content. *The Review of Financial Studies*, 18(4), 1305-1342.
- Jorion, P., & Goetzmann, W. N. (1999). Global stock markets in the twentieth century. *The journal of finance*, 54(3), 953-980.
- Joseph, K., Wintoki, M. B., & Zhang, Z. (2011). Forecasting abnormal stock returns and trading volume using investor sentiment: Evidence from online search. *International Journal of Forecasting*, 27(4), 1116-1127.
- Kahneman, D. (1973). Attention and effort (Vol. 1063, pp. 218-226). Englewood Cliffs, NJ: Prentice-Hall.

- Kahneman, D. and A. Tversky. (1979). "Prospect Theory: An Analysis of Decision Under Risk." *Econometrica*, 47:263-291
- Karemera, D., Ojah, K., & Cole, J. A. (1999). Random walks and market efficiency tests: Evidence from emerging equity markets. *Review of Quantitative finance and Accounting*, 13, 171-188.
- Karkowska, R., & Palczewski, A. (2023). Does high-frequency trading actually improve market liquidity? A comparative study for selected models and measures. *Research in International Business and Finance*, 64, 101872.
- Karpoff, J. M. (1986). A theory of trading volume. *The journal of finance*, 41(5), 1069-1087.
- Khan, M. S., & Fahim, M. M. U. (2021). The four-factor model and stock returns in Bangladesh. *International Journal of Accounting & Finance Review*, 6(2), 133-149.
- Kim N, Lučivjanská K, Molnár P, Villa R (2019) Google searches and stock market activity: evidence from Norway. *Financ Res Lett* 28:2 08–220. <https://doi.org/10.1016/j.frl.2018.05.003>
- Koch, S., & Dimpfl, T. (2023). Attention and retail investor herding in cryptocurrency markets. *Finance Research Letters*, 51, 103474.
- Koester, A., Lundholm, R., & Soliman, M. (2016). Attracting attention in a limited attention world: Exploring the causes and consequences of extreme positive earnings surprises. *Management Science*, 62(10), 2871-2896.
- Konstantinidis, A., Katarachia, A., Borovas, G., & Voutsas, M. E. (2012). From efficient market hypothesis to behavioural finance: Can behavioural finance be the new dominant model for investing. *Scientific Bulletin–Economic Sciences*, 11(2), 16-26.
- Kristoufek L (2013) Can Google Trends search queries contribute to risk diversification? *Sci Rep* 3(1):1–5. <https://doi.org/10.1038/srep02713>
- Kristoufek L (2015) Power-law correlations in finance-related Google searches, and their cross-correlations with volatility and traded volume: evidence from the Dow Jones Industrial components. *Physica A* 428:194–205. <https://doi.org/10.1016/j.physa.2015.02.057>

- Kumari, V., Kumar, S., Pandey, D. K., & Gupta, P. (2024). Dividend announcements and stock returns: a retrospective analysis. *Qualitative Research in Financial Markets*, 16(4), 709-740.
- Laboissiere LA, Fernandes RA, Lage GG (2015) Maximum and minimum stock price forecasting of Brazilian power distribution companies based on artificial neural networks. *Appl Soft Comput* 35:66–74. <https://doi.org/10.1016/j.asoc.2015.06.005>
- Lahmiri, S. (2017). Modeling and predicting historical volatility in exchange rate markets. *Physica A: Statistical Mechanics and its Applications*, 471, 387-395.
- Lai, H. H., Chang, T. P., Hu, C. H., & Chou, P. C. (2022). Can google search volume index predict the returns and trading volumes of stocks in a retail investor dominant market. *Cogent Economics & Finance*, 10(1), 2014640.
- Lambert, R. A., Leuz, C., & Verrecchia, R. E. (2012). Information asymmetry, information precision, and the cost of capital. *Review of finance*, 16(1), 1-29.
- Levin, I. P., S. L. Schneider, and G. J. Gaeth. "All Frames Are Not Created Equal: A Typology and Critical Analysis of Framing Effects." *Organizational Behaviour and Human Decision Processes*, 76, (1998), pp. 149–188
- Li, C., Huang, W., Wang, W. S., & Chia, W. M. (2021). Price change and trading volume: Behavioural heterogeneity in stock market. *Computational Economics*, 1-37.
- Li, X., Hendler, J. A., & Teall, J. L. (2016). Investor attention on the social web. *Journal of Behavioural Finance*, 17(1), 45-59.
- Li, X., Mahani, R. S., & Sandhya, V. (2011). Does investor attention affect stock prices?. *Available at SSRN 1748851*.
- Lin, M. C., Wu, C. H., & Chiang, M. T. (2014). Investor attention and information diffusion from analyst coverage. *International Review of Financial Analysis*, 34, 235-246.
- Lintner, J. (1965) 'The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets', *Review of Economics and Statistics*, Vol. 47, No. 1, pp.13–37.
- Liu, S. (2015). Investor sentiment and stock market liquidity. *Journal of Behavioural Finance*, 16(1), 51-67.

- Lo, A. W., & Wang, J. (2010). Stock market trading volume. In *Handbook of financial econometrics: Applications* (pp. 241-342). Elsevier.
- Lobão J, Pacheco L, Pereira C (2017) The use of the recognition heuristic as an investment strategy in European stock markets. *J Econ Finance Adm Sci* 22(43):207–223. <https://doi.org/10.1108/JEFAS-01-2017-0013>
- London Stock Exchange, LSEG Workspace, 26 June 2024
- Lou, D. (2014). Attracting investor attention through advertising. *The Review of Financial Studies*, 27(6), 1797-1829.
- Lundblad, C. (2007). The risk return tradeoff in the long run: 1836–2003. *Journal of Financial Economics*, 85(1), 123-150.
- Lyócsa Š, Baumöhl E, Výrost T, Molnár P (2020) Fear of the coronavirus and the stock markets. *Finance Res Lett* 36:101735
- Malceniece, L., Malcenieks, K., & Putniņš, T. J. (2019). High frequency trading and comovement in financial markets. *Journal of Financial Economics*, 134(2), 381-399.
- Markowitz, H. M. (1991). Foundations of portfolio theory. *The journal of finance*, 46(2), 469-477.
- Merton, R. C. (1987). A simple model of capital market equilibrium with incomplete information.
- Mitchell, M. L., & Mulherin, J. H. (1994). The impact of public information on the stock market. *The Journal of Finance*, 49(3), 923-950.
- Molnár, P. (2012). Properties of range-based volatility estimators. *International Review of Financial Analysis*, 23, 20-29.
- Mondria J, Wu T, Zhang Y (2010) The determinants of international investment and attention allocation: using internet search query data. *J Int Econ* 82(1):85–95. <https://doi.org/10.1016/j.jinteco.2010.04.007>
- Moussa F, BenOuda O, Delhoumi E (2017a) The use of open source internet to analysis and predict stock market trading volume. *Res Int Bus Finance* 41:399–411. <https://doi.org/10.1016/j.ribaf.2017.04.048>
- Moussa F, Delhoumi E, Ouda OB (2017b) Stock return and volatility reactions to information demand and supply. *Res Int Bus Finance* 39:54–67. <https://doi.org/10.1016/j.ribaf.2016.07.016>

- Muth, J. F. (1981). Rational expectations and the theory of price movements. *Econometrica*, 29(3), 3.
- Mytton, O. T., Rutter, P. D., & Donaldson, L. J. (2012). Influenza A (H1N1) pdm09 in England, 2009 to 2011: a greater burden of severe illness in the year after the pandemic than in the pandemic year. *Eurosurveillance*, 17(14), 20139.
- Nazário, R. T. F., e Silva, J. L., Sobreiro, V. A., & Kimura, H. (2017). A literature review of technical analysis on stock markets. *The Quarterly Review of Economics and Finance*, 66, 115-126.
- Nguyen C, Hoang L, Shim J, Truong P (2020) Internet search intensity, liquidity and returns in emerging markets. *Res Int Bus Finance* 52:101166. <https://doi.org/10.1016/j.ribaf.2019.101166>
- Nguyen CP, Schinckus C, Nguyen TVH (2019) Google search and stock returns in emerging markets. *Borsa Istanbul Rev* 9(4):288–296. <https://doi.org/10.1016/j.bir.2019.07.001>
- Nyberg, H. (2012). Risk-return tradeoff in US stock returns over the business cycle. *Journal of Financial and Quantitative Analysis*, 47(1), 137-158.
- Oliveira-Brochado A (2019) Google search-based sentiment indexes. *IIMB Manag Rev*. <https://doi.org/10.1016/j.iimb.2019.10.015>
- Osabuohien-Irabor, O. (2021). Investors' attention: does it impact the Nigerian stock market activities?. *Journal of Economics and Development*, 23(1), 59-76.
- Osabuohien-Irabor, O. (2021). Investors' attention: does it impact the Nigerian stock market activities? *Journal of Economics and Development*, 23(1), 59-76
- Owusu, S. P., & Laryea, E. (2023). The impact of anchoring bias on investment decision-making: evidence from Ghana. *Review of Behavioural Finance*, 15(5), 729-749.
- 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), 30.
- Papadamou S, Koulis A, Kyriakopoulos C, Fassas AP (2022) Cannabis stocks returns: the role of liquidity and investors' attention via google metrics. *International Journal of Financial Studies* 10(1):7. <https://doi.org/10.3390/ijfs10010007>

- Park, C. H., & Irwin, S. H. (2007). What do we know about the profitability of technical analysis?. *Journal of Economic surveys*, 21(4), 786-826.
- Pati, P. C., Barai, P., & Rajib, P. (2018). Forecasting stock market volatility and information content of implied volatility index. *Applied Economics*, 50(23), 2552-2568.
- Peng, L., & Xiong, W. (2006). Investor attention, overconfidence and category learning. *Journal of Financial Economics*, 80(3), 563-602.
- Pereira EJDAL, Da Silva MF, Da Cunha Lima IC, Pereira HBB (2018) Trump's effect on stock markets: a multiscale approach. *Physica A* 512:241–247. <https://doi.org/10.1016/j.physa.2018.08.069>
- Perlin, M. S., Caldeira, J. F., Santos, A. A., & Pontuschka, M. (2017). Can we predict the financial markets based on Google's search queries? *Journal of Forecasting*, 36(4), 454-467.
- Petropoulos, A., Siakoulis, V., Stavroulakis, E., Lazaris, P., & Vlachogiannakis, N. (2022). Employing google trends and deep learning in forecasting financial market turbulence. *Journal of Behavioural Finance*, 23(3), 353-365.
- Pham, L., & Huynh, T. L. D. (2020). How does investor attention influence the green bond market?. *Finance Research Letters*, 35, 101533.
- Piñeiro-Chousa J, López-Cabarcos MÁ, Ribeiro-Soriano D (2020) Does investor attention influence water companies' stock returns? *Technol Forecast Soc Chang* 158:120115. <https://doi.org/10.1016/j.techfore.2020.120115>
- Poon, S. H., & Granger, C. W. J. (2003). Forecasting volatility in financial markets: A review. *Journal of economic literature*, 41(2), 478-539.
- Preis T, Moat HS, Stanley HE (2013) Quantifying trading behaviour in financial markets using Google Trends. *Sci Rep* 3(1):1–6. <https://doi.org/10.1038/srep01684>
- Preis T, Reith D, Stanley HE (2010) Complex dynamics of our economic life on different scales: insights from search engine query data. *Philos Trans R Soc A Math Phys Eng Sci* 368(1933):5707–5719. <https://doi.org/10.1098/rsta.2010.0284>
- Pyun, C. (2024). The Wikipedia effect: Analyzing investor attention for strategic investment decisions. *Economics Letters*, 241, 111836.

- Ramos SB, Latoeiro P, Veiga H (2020) Limited attention, salience of information and stock market activity. *Econ Model* 87:92–108. <https://doi.org/10.1016/j.econmod.2019.07.010>
- Rangapriya, S., & Lokhande, M. (2023). Testing for Market Efficiency Using News-Driven Sentiment: Evidence from Select NYSE Stocks. In *Artificial Intelligence for Capital Markets* (pp. 128-140). Chapman and Hall/CRC.
- Reber, R., & Schwarz, N. (1999). Effects of perceptual fluency on judgments of truth. *Consciousness and cognition*, 8(3), 338-342.
- Salisu AA, Vo XV (2021) Firm-specific news and the predictability of consumer stocks in Vietnam. *Finance Res Lett* 41:101801. <https://doi.org/10.1016/j.frl.2020.101801>
- Saxena, K., & Chakraborty, M. (2022). Does it pay to pay attention to attention? Evidence from an emerging market. *Managerial Finance*, 48(4), 629-642.
- Seabold, S., & Coppola, A. (2015). Nowcasting prices using Google trends: an application to Central America. *World Bank Policy Research Working Paper*, (7398).
- Seasholes, M. S., & Wu, G. (2007). Predictable behaviour, profits, and attention. *Journal of Empirical Finance*, 14(5), 590-610.
- Seifter, A., Schwarzwald, A., Geis, K., & Aucott, J. (2010). The utility of Google Trends for epidemiological research: Lyme disease as an example. *Geospatial health*, 4(2), 135-137.
- Sen, A. (1997). Maximization and the Act of Choice. *Econometrica: Journal of the Econometric Society*, 745-779.
- Sharpe, W. F. (1965). Risk-aversion in the stock market: Some empirical evidence. *The Journal of Finance*, 20(3), 416-422.
- Sharpe, W.F. (1964) 'Capital asset prices: a theory of market equilibrium under conditions of risk', *Journal of Finance*, Vol. 19, No. 3, pp.425–442.
- Shefrin, H. (2002). *Beyond greed and fear: Understanding behavioural finance and the psychology of investing*. Oxford University Press.
- Shen D, Wang C (2023) A systematic review of investor attention: measurements, implications, and future directions. *Artif Intell Learn Comput Econ Finance*. https://doi.org/10.1007/978-3-031-15294-8_8

- Sifat IM, Thaker HMT (2020) Predictive power of web search behaviour in five ASEAN stock markets. *Res Int Bus Finance* 52:101191. <https://doi.org/10.1016/j.ribaf.2020.101191>
- Sims, C. A. (2003). Implications of rational inattention. *Journal of monetary Economics*, 50(3), 665-690.
- Škrinjarić, T. (2019). Time varying spillovers between the online search volume and stock returns: Case of CESEE markets. *International journal of financial studies*, 7(4), 59.
- Smales LA (2021) Investor attention and global market returns during the COVID-19 crisis. *Int Rev Financ Anal* 73:101616.
- So, S. M., & Lei, V. U. (2015). On the relationship between investor sentiment, VIX and trading volume. *Risk Governance & Control: Financial Markets & Institutions*, 5(4), 114-122.
- StatCounter. (February 12, 2024). Market share of leading desktop search engines worldwide from January 2015 to January 2024 [Graph]. In Statista. Retrieved July 26, 2024, from <https://www.statista.com/statistics/216573/worldwide-market-share-of-search-engines/>
- StatCounter. (July 8, 2024). Market share of leading search engines in China from June 2019 to June 2024 [Graph]. In Statista. Retrieved July 26, 2024, from <https://www.statista.com/statistics/253340/market-share-of-search-engines-in-china-pageviews/>
- Statman, M. (1995, December). Behavioural finance versus standard finance. In AIMR conference Proceedings (Vol. 7, pp. 14-22). CFA Institute.
- Statman, M. (1999). Behavioural finance: Past battles and future engagements. *Financial analysts journal*, 55(6), 18-27
- Statman, M., Thorley, S., & Vorkink, K. (2006). Investor overconfidence and trading volume. *The Review of Financial Studies*, 19(4), 1531-1565.
- Suardi, S., Rasel, A. R., & Liu, B. (2022). On the predictive power of tweet sentiments and attention on bitcoin. *International Review of Economics & Finance*, 79, 289-301.

- Subramaniam S, Chakraborty M (2021) COVID-19 fear index: Does it matter for stock market returns? *Review of Behavioural Finance* 13(1):40–50. <https://doi.org/10.1108/RBF-08-2020-0215>
- Sukamulja, S., & Thedora, S. (2023). Could Investors' Attention in News Hunting Affect the International Capital Market?. *The Indonesian Capital Market Review*, 15(2), 3.
- Swamy V, Dharani M (2019) Investor attention using the Google search volume index: impact on stock returns. *Rev Behav Finance* 11(1): 55–69. <https://doi.org/10.1108/RBF-04-2018-0033>
- Tajmazinani M, Hassani H, Raei R, Rouhani S (2022) Modeling stock price movements prediction based on news sentiment analysis and deep learning. *Ann Financ Econ (AFE)* 17(01):1–19.
- Takeda F, Wakao T (2014) Google search intensity and its relationship with returns and trading volume of Japanese stocks. *Pac Basin Finance J* 27:1–18. <https://doi.org/10.1016/j.pacfin.2014.01.003>
- Tan SD, Taş O (2019) Investor attention and stock returns: evidence from Borsa Istanbul. *Borsa Istanbul Rev* 19(2):106–116. <https://doi.org/10.1016/j.bir.2018.10.003>
- Tang W, Zhu L (2017) How security prices respond to a surge in investor attention: evidence from Google Search of ADRs. *Glob Financ J* 33:38–50. <https://doi.org/10.1016/j.gf.2016.09.001>
- Tantaopas P, Padungsaksawasdi C, Treepongkaruna S (2016) Attention effect via internet search intensity in Asia-Pacific stock markets. *Pac Basin Finance J* 38:107–124
- Tazi, O., Aguenau, S., & Abrache, J. (2022). A comparative study of the Fama-French three factor and the Carhart four factor models: Empirical Evidence from Morocco. *International Journal of Economics and Financial Issues*, 12(1), 58-66.
- Thaler, R. H. (2016). Behavioural economics: Past, present, and future. *American economic review*, 106(7), 1577-1600.
- Tkac, P. A. (1999). A trading volume benchmark: Theory and evidence. *Journal of Financial and Quantitative Analysis*, 34(1), 89-114.

- Varamini, H., & Kalash, S. (2008). Testing market efficiency for different market capitalization funds. *American Journal of Business*, 23(2), 17-28.
- Vilutiene, L., & Dumciuviene, D. (2020). The impact of international capital flows on Central and Eastern European Countries' investments, savings, consumption, and current accounts. *Engineering Economics*, 31(4), 450-460.
- Vlastakis, N., & Markellos, R. N. (2012). Information demand and stock market volatility. *Journal of Banking & Finance*, 36(6), 1808-1821.
- Vo, D. H. (2023). Volatility spillovers across sectors and their magnitude: A sector-based analysis for Australia. *Plos one*, 18(6), e0286528.
- Vosen, S., & Schmidt, T. (2011). Forecasting private consumption: survey-based indicators vs. Google trends. *Journal of forecasting*, 30(6), 565-578.
- Vozlyublennaiia N (2014) Investor attention, index performance, and return predictability. *J Bank Finance* 41:17–35. <https://doi.org/10.1016/j.jbankfin.2013.12.010>
- Wäneryd, K. E. (2001). Stock-market psychology: How people value and trade stocks. Edward Elgar Publishing.
- Woloszko, N. (2020). Tracking activity in real time with Google Trends.
- World Bank. "Purchasing power parities and the size of world economies: Results from the 2017 International Comparison Program." (2020).
- Wright C, Swidler S (2023) Abnormal trading volume, news and market efficiency: evidence from the Jamaica stock exchange. *Res Int Bus Financ* 64:101804. <https://doi.org/10.1016/j.ribaf.2022.101804>
- Wylie, G., & Allport, A. (2000). Task switching and the measurement of "switch costs". *Psychological research*, 63, 212-233.
- Xiao, J., & Wang, Y. (2021). Investor attention and oil market volatility: Does economic policy uncertainty matter?. *Energy Economics*, 97, 105180
- Yamamoto, R. (2012). Intraday technical analysis of individual stocks on the Tokyo Stock Exchange. *Journal of Banking & Finance*, 36(11), 3033-3047.
- Yan, X., & Zheng, L. (2017). Fundamental analysis and the cross-section of stock returns: A data-mining approach. *The Review of Financial Studies*, 30(4), 1382-1423.

- Yandex. (February 6, 2024). Distribution of search engine traffic in Russia from 3rd quarter 2019 to 4th quarter 2023, by company [Graph]. In Statista. Retrieved July 26, 2024, from <https://www.statista.com/statistics/1094920/leading-search-engines-by-visits-share-russia/>
- Yang, A. C., Huang, N. E., Peng, C. K., & Tsai, S. J. (2010). Do seasons have an influence on the incidence of depression? The use of an internet search engine query data as a proxy of human affect. *PloS one*, 5(10), e13728.
- Yang, D., Ma, T., Wang, Y., & Wang, G. (2021). Does investor attention affect stock trading and returns? Evidence from publicly listed firms in China. *Journal of Behavioural Finance*, 22(4), 368-381.
- Yin, S., & Ho, M. (2012). Monitoring a toxicological outbreak using Internet search query data. *Clinical toxicology*, 50(9), 818-822.
- Yoshinaga C, Rocco F (2020) Investor attention: can google search volumes predict stock returns? *BBR Braz Bus Rev* 17:523–539. <https://doi.org/10.15728/bbr.2020.17.5.3>
- Yuan, Y. (2015). Market-wide attention, trading, and stock returns. *Journal of financial economics*, 116(3), 548-564.
- Zhang, W., Shen, D., Zhang, Y., & Xiong, X. (2013). Open source information, investor attention, and asset pricing. *Economic Modelling*, 33, 613-619.
- Zhang, W., Yan, K., & Shen, D. (2021). Can the Baidu Index predict realized volatility in the Chinese stock market?. *Financial Innovation*, 7, 1-31.
- Zhang, X., Huang, Y., Xu, K., & Xing, L. (2023). Novel modelling strategies for high-frequency stock trading data. *Financial Innovation*, 9(1), 39
- Zhang, X., Lu, F., Tao, R., & Wang, S. (2021). The time-varying causal relationship between the Bitcoin market and internet attention. *Financial Innovation*, 7, 1-19.
- Zhao, W., & Zhang, J. (2024). Investor Attention and Stock Liquidity in the Chinese Market. *International Advances in Economic Research*, 30(1), 65-82.

X – APPENDIXES

X - APPENDIXES

APPENDIX 1. List of companies included in the data sample.

Company	Country	Industry	Ticker	Keyword
Adyen	Netherland	Technology	ADYEN.AS	ADYEN
Ahold Delhaize	Netherland	Consumer Non-Cyclicals	AD.AS	AD
Air Liquide	France	Basic materials	AIRP.PA	AIRP
Airbus Group	Netherland	Industrials	AIR.PA	AIR
Amadeus IT Group	Spain	Technology	AMA.MC	AMA
Anheuser Busch Inbev	Belgium	Consumer Non-Cyclicals	ABI.BR	ABI
ASML Holding	Netherland	Technology	ASML.AS	ASML
Barclays	UK	Financials	BARC.L	BARC
BBVA	Spain	Financials	BBVA.MC	BBVA
BG Grp	UK	Energy	BG.L^B16	BG
BNP Paribas	France	Financials	BNPP.PA	BNPP
BT Group	UK	Technology	BT.L	BT

Compagnie Financiere Richemont	Switzerland	Consumer Cyclical	CFR.S	CFR
CRH	Ireland	Industrials	CRH.I^I23	CRH
Danone	France	Consumer Non-Cyclicals	DANO.PA	DANO
Enel	Italy	Utilities	ENEI.MI	ENEI
Engie	France	Utilities	ENGIE.PA	ENGIE
Eni SpA	Italy	Energy	ENI.MI	ENI
Ferrari	Italy	Consumer Cyclical	RACE.MI	RACE
Fresenius	Germany	Healthcare	FREG.DE	FREG
Glencore	UK	Basic materials	GLEN.L	GLEN
Hermes International	France	Consumer Cyclical	HRMS.PA	HRMS
Iberdrola	Spain	Utilities	IBE.MC	IBE
Imperial Brands	UK	Consumer Non-Cyclicals	IMB.L	IMB
Inditex	Spain	Consumer Cyclical	ITX.MC	ITX
ING Groep	Netherlands	Financials	INGA.AS	INGA
Intesa Sanpaolo	Italy	Financials	ISP.MI	ISP
Linde	Germany	Basic materials	LIN.DE	LIN

Lloyds Bank UK	UK	Financials	LLOY.L	LLOY
Louis Vuitton	France	Consumer Cyclicals	LVMH.PA	LVMH
National Grid	UK	Utilities	NG.L	NG
NOKIA	Finland	Technology	NOKIA.HE	NOKIA
Orange	France	Technology	ORAN.PA	ORAN
Pernod Ricard	France	Consumer Non-Cyclicals	PERP.PA	PERP
Philips	Netherlands	Healthcare	PHG.AS	PHG
Prosus	Netherlands	Technology	PRX.AS	PRX
Prudential	UK	Financials	PRU.L	PRU
Relx	UK	Technology	REL.L	REL
Safran	France	Industrials	SAF.PA	SAF
Sanofi	France	Healthcare	SASY.PA	SASY
Santander	Spain	Financials	SAN.MC	SAN
Schneider Electric	France	Industrials	SCHN.DE	SCHN
Société Générale	France	Financials	SOGN.PA	SOGN
StanChart	UK	Financials	STAN.L	STAN
Telefónica	Spain	Technology	TEF.MC	TEF
Tesco	UK	Consumer Non-Cyclicals	TSCO.L	TSCO

Uni Credit		Italy	Financials	CRDI.MI	CRDI
Unibail Rod West		France	Financials	URW.PA	URW
Unilever		UK	Consumer Non-Cyclicals	UNA.AS^K20	UNA
Universal	Music	Netherland	Consumer Cyclicals	UMG.AS	UMG
Group					
Vivendi		France	Consumer Cyclicals	VIV.PA	VIV
Vodafone		UK	Technology	VOD.L	VOD

Source: Own elaboration.

This blank sheet is the **back cover**, so there is no heading (there is a change of section).