



SCHOOL OF ECONOMICS AND MANAGEMENT

Department of Business Administration

BUSN79 - Degree Project in Accounting and Finance

28th May 2023

A Catering Theory of Investment Policies

Do U.S. firms cater to sentiment?

Authors:

Philip Lindquist

Nicolas Nolén

Supervisor

Elias Bengtsson

Summary

Seminar date: 2023-06-01

Course: BUSN79 - Degree Project in Accounting and Finance

Authors: Philip Lindquist & Nicolas Nolén

Advisor/Examiner: Elias Bengtsson/Reda Moursli

Five key words: “Retail Investor Sentiment”, “FEARS Index”, “Corporate Investment Policies”, “Behavioral Corporate Finance”, “VIX”

Purpose: To investigate retail investors’ growing influence in shaping corporate investment policies. It further seeks to understand whether firms cater to retail investor sentiment.

Methodology: We employ a multiple linear regression model of a strongly-balanced panel data set, incorporating fixed effects and clustered standard errors. We construct a Financial and Economic Attitudes Revealed by Search (“FEARS”) index by aggregating U.S. households' Google search volume data to measure retail investor sentiment.

Theoretical perspectives: To support our analysis, we review various relevant theories, including market timing, catering, prospect, and herd behavior theories.

Empirical foundation: Our sample data was retrieved from Bloomberg. It consists of 23,712 quarterly observations on 741 publicly listed firms on the Nasdaq Composite.

Conclusions: In a subsample analysis, we find a significant positive relationship between retail investor sentiment and investment activities before and after the onset of the pandemic. Further, we find an inverse relationship between the FEARS index and VIX, suggesting that retail and institutional investors may influence corporate investment activities differently.

Table of contents

1. Introduction.....	4
1.1 Background.....	4
1.2 Problematization.....	5
1.3 Purpose & Research Question.....	6
1.4 Empirical Findings.....	6
1.5 Contribution.....	7
1.6 Structure of the paper.....	7
2. Theoretical review.....	7
2.1 Investment Policies.....	7
2.2 Rational Managers in Irrational Market.....	9
2.3 The Prospect Theory & Managerial Bias.....	10
2.4 The Role of Cognitive Biases & Herding Behavior Theory.....	11
2.5 Critique & Limitations of the Theoretical Framework.....	12
3. Empirical Review.....	13
3.1 Unveiling the Influence of Catering & Market Timing on Corporate Investments.....	13
3.2 The Hidden Power of Managerial Biases on Corporate Investments.....	15
3.3 How Investor Sentiment Shapes Corporate Investments.....	16
3.4 Leveraging Textual Analysis to Gauge Investor Sentiment.....	17
4. Literature Summary & Hypotheses Development.....	19
5. Data & Sample Description.....	21
5.1 Sample Description.....	21
5.2 Variable definition.....	23
5.2.1 Dependent Variable.....	23
5.2.2. Independent variables.....	24
5.3 Summary Statistics.....	27
6. Methodology.....	29
6.1 The FEARS Index Proxy.....	29
6.1.1 Google Trends.....	30
6.1.2 Collecting & Adjusting the Google Trends Data.....	30
6.1.3 Constructing the Index.....	31
6.2 Univariate analysis.....	33
6.3 Pre-regression Diagnostics.....	33
6.4 Model specification.....	34
6.5 Subsample Analysis.....	35
7. Empirical Findings & Analysis.....	36
7.1 Univariate Analysis.....	36
7.2 Multivariate Analysis.....	37
7.2.1 Hypothesis 1.....	37

7.2.2 Hypothesis 2.....	38
7.2.3 Hypothesis 3.....	41
7.3 Robustness Tests.....	42
8. Reflection & Discussion.....	43
8.1. Model limitations.....	43
8.1.1 Discussion of the FEARS-index as a proxy for retail investor sentiment.....	44
8.2 Discussion in relation to H1.....	44
8.3 Discussion in relation to H2 and endogeneity in the model.....	45
9. Conclusion.....	46
Reference list.....	48
Figures.....	57
Tables.....	60
Appendices.....	70

1. Introduction

1.1 Background

According to a study by Bloomberg Intelligence, 19.5% of all stock market shares traded during the first half of 2020 were attributed to retail investors. This figure is nearly double the volume observed in 2010, reflecting an accelerated shift in market dynamics over the past decade (Forbes, 2022). This surge in retail investor activity emphasizes a new financial landscape where the influence and power of retail investors in financial markets are growing, fuelled by the growth of digital trading platforms (Wall Street Journal, 2020). This has democratized and made the stock market available for millions of new investors, a domain previously dominated by institutional investors (Deloitte, 2021). Significant events such as the GameStop short squeeze reflect this transformative shift. Such events subsequently prompted the International Organization of Securities Commissions to review the suitability of existing regulatory frameworks considering the influence digital tools and social media exerts on investors' behavior, acknowledging the possibility of emotional decision-making surpassing the significance of fundamental analysis (IOSCO, 2022).

The transformed market dynamics have brought renewed relevance to Keynes's (1937) concept of "animal spirits" to fully comprehend this phenomenon, illustrating how investors' irrational behavior can affect market dynamics. Keynes's (1937) insights set the stage for an evolving discourse on the influence of investor sentiment in financial markets, compelling us to consider a critical question: As the sphere of retail investor activity continues to grow, how can this impact effectively be quantified?

We propose the utilization of the Financial and Economic Attitudes Revealed by Search ("FEARS") index by Da, Engelberg, and Gao (2015) as a potential proxy for retail investor sentiment. This search-based measure, derived from aggregated Google Search queries, offers a more transparent reflection of retail sentiment than traditional measures that often fail to differentiate between retail and institutional investor sentiment. Previous research has established a positive correlation between market-based proxies of investor sentiment and corporate

investments (Baker & Wurgler, 2007; Polk & Sapienza, 2009; Zhu, Song & Huang, 2018; Du & Hu, 2019; Li & Grundy, 2023). Retail investors' increasing influence and power prompt the inquiry of whether they can influence corporate investment decisions and whether firms negotiate a balance between maximizing shareholder value and catering to these “animal spirits”.

1.2 Problematicization

Traditionally, researchers have predominantly adopted two approaches to gauge investor sentiment. The first approach employs market-based proxies such as option implied volatilities (“VIX”) and mutual fund flows. Although these market-based indicators offer high-frequency data and are readily available, they are the equilibrium outcome of many economic forces other than investor sentiment (Da et al., 2015). Thus, it is challenging to discern exactly what these market-based measures capture due to potential correlations with omitted variables or other unmeasured factors.

The second approach turns to survey-based indices. While survey-based measures seek to assess investor sentiment directly, their efficacy depends heavily on participant honesty and engagement. With little incentive to answer carefully or truthfully, especially for sensitive questions, survey-based measures may not accurately capture the true sentiment (Singer, 2002). High non-response rates and low incentives for truth-telling further undermine their reliability. Thus, as search-based measures reveal attitudes rather than inquire about them, they might present a more authentic picture of sentiment.

Therefore, the main limitation of previous studies stems from the inherent shortcoming of traditional investor sentiment measures. These limitations potentially lead to an incomplete or distorted understanding of the influence of investor sentiment on corporate investments. Our thesis aims to address this gap by utilizing the FEARS index, a search-based proxy for retail investor sentiment, to offer a more comprehensive and nuanced examination of how investor sentiment impacts corporate investment decisions. Previous methodologies may have been flawed due to their inability to isolate investor sentiment using market-based measures and the

limited incentives for truth-telling in survey-based measures. Consequently, the FEARS index thus holds promise as a more precise and reliable alternative.

1.3 Purpose & Research Question

The purpose of this study is twofold. Firstly, it seeks to investigate retail investors' growing influence and power in shaping corporate investment decisions. With the increasing participation of retail investors in the stock market, it is pertinent to explore whether their sentiment, as proxied by the FEARS index, has a consequential impact on corporate investment policies. Secondly, the study aims to understand whether firms engage in a balancing act between adhering to traditional financial principles aimed at maximizing shareholder value and catering to the market sentiment, often swayed by retail investors' collective “animal spirits”. Consequently, this paper aims to answer the following research questions:

RQ1: *Does retail investor sentiment, as proxied by the FEARS index, affect U.S. firms' investment decisions?*

RQ2: *What relationship, if any, exists between the pursuit of maximizing shareholder value and catering to market sentiment in corporate investment decisions?*

1.4 Empirical Findings

This paper utilizes a strongly-balanced panel of 741 publicly listed firms on the Nasdaq Composite from 2014 to 2021. Employing multiple linear regression models with fixed effects and clustered standard errors, our initial analysis did not reveal a statistically significant relationship between corporate investment policies and retail investor sentiment. However, in a pre-Covid and Covid subsample analysis, we find a robust correlation between retail investor sentiment and investment activities, but we are unable to infer the nature of the relationship. Thus, we cannot conclude if retail investor sentiment affects investment decisions or if firms cater to retail investors through investment activities. Further, this study finds an inverse relationship between the FEARS index and VIX, suggesting that retail and institutional investors may influence corporate investment activities differently.

1.5 Contribution

This study takes an innovative approach by utilizing the FEARS index, a search-based measure, to investigate the effects of investor sentiment on firms' investment activities in the U.S. Thus, this method overcomes the limitations of conventional market-based measures, offering a more transparent and accurate representation of investor sentiment. By undertaking this approach, we further widen the spectrum of sentiment measurements, isolating retail investors. This aspect of our study is especially relevant considering the growing influence of retail investors in financial markets, a dimension that has received limited attention in earlier research, which has mainly been focused on institutional investors. We further examine the dynamics of investor sentiment by considering both positive and negative sentiment, thereby expanding the knowledge of its reciprocal influence on corporate investments. This particular aspect of our study contributes depth and nuance to the existing literature, which primarily focuses on the impact of negative sentiment (Tetlock, 2007; Tetlock, Saar-Tsechansky & Macskassy, 2008; Da et al., 2015).

1.6 Structure of the paper

The structure of the paper proceeds as follows: Section 2 explores the theoretical framework and its critiques. Section 3 surveys the empirical literature. Our hypotheses, derived from the literature, are presented in Section 4. Section 5 describes our data and sample, followed by the methodological design in Section 6. We share our empirical findings and analysis in Section 7 and briefly discuss the study in Section 8. Section 9 concludes the study.

2. Theoretical review

2.1 Investment Policies

Traditional theories in finance are built on the premise of rational agents that exhibit utility-maximizing behavior. For instance, the Efficient Market Hypothesis ("EMH") presumes that financial markets are perfectly rational and that all information is completely and immediately reflected in asset prices (Fama, 1970). In perfect capital markets, Modigliani and Miller (1958) suggest that the value of any firm is solely determined by its investment decisions and free cash flows rather than its optimal capital structure. In other words, they fundamentally

assert that firm value is exclusively guided by its real assets, namely investment policies, which are irrelevant to its financing policies. Hence, the optimal investment policy, according to Modigliani and Miller (1958), is to invest in every project featuring a positive net present value (“NPV”). Since the introduction of the Modigliani-Miller theorem, subsequent researchers and academics have expanded this and developed their theories in a setting of market frictions.

The pecking order theory by Myers and Majluf (1984) suggests that firms prefer different sources of financing to finance their investments based on information asymmetries, which emerge due to adverse selection issues arising from management-stakeholder discrepancies. However, the trade-off theory by Myers (1984) assumes that an optimal capital structure is achieved at the equilibrium where the tax benefits of debt are balanced against the costs of financial distress, ultimately guiding firms to choose the capital structure that minimizes the cost of capital. The costs of financial distress are a fundamental aspect of the trade-off theory since they might negatively affect a firm's reputation and stakeholder relationships, which could negatively affect its investment policies as it can limit its access to capital and affect the cost of capital. Likewise, an extension of the market timing theory by Baker and Wurgler (2002) proposes that managers might assess the current market conditions and adjust their investment and financing policies accordingly to minimize the cost of capital, hence encouraging investments.

A crucial similarity these theories share is that firms strive to achieve the highest possible NPV when making investment decisions. A related concept within this context is the internal rate of return (“IRR”), the discount rate that sets the NPV of all cash flows from a project equal to zero. According to Cheng, Kite, and Radtke (1994), the NPV and IRR are affected by numerous parameters, including cash flows and cost of capital. As such, these two methods are crucial for firms to evaluate different projects, thereby guiding firms’ investment decisions.

Expanding upon the cost of capital, Baker and Wurgler (2013) argue that managers' tendencies to overinvest are mitigated when equity is issued. This may be explained by managers’ inclination to preferably issue debt, primarily due to higher costs of capital and ownership dilution

associated with equity. They further assimilate the need for equity with limited cash or cash flows, implying that firms' investment levels strongly correlate with their current cash flows and earnings. Consequently, Baker and Wurgler (2013) suggest that these assumptions and ideas pave the way to a behavioral base for Jensen's (1986) hypothesis on agency costs of free cash flow. However, in contrast to Jensen (1986), they assert that these agency costs are not due to managerial entrenchment but rather overly-optimistic managers overinvesting.

Jensen's (1986) free cash flow hypothesis posits that excessive free cash flows, exceeding what is necessary to meet debt obligations and undertake all positive NPV projects, can lead to agency problems as managers may engage in wasteful spending or overinvestment. This is known as a managerial entrenchment problem, where the interests of managers and shareholders are misaligned. To mitigate this problem, Jensen (1986) proposed that firms increase their leverage or disburse excess cash to their shareholders as dividends, restricting the pool of free cash flows accessible for managerial discretion.

2.2 Rational Managers in Irrational Market

The behavioral explanation of Jensen's (1986) hypothesis and the catering theory of managers' behavior might offer complementary perspectives on the dynamics that may influence corporate investment decisions. In 2004, Baker and Wurgler introduced the catering theory, suggesting that firms have catering incentives to shape their dividend policies based on investors' needs and preferences. They assume that investors' demands and preferences are time-varying due to the dynamic nature of market sentiment and macroeconomic conditions. Baker and Wurgler (2004) further suggest that managers act in the best interests of their shareholders, which may incentivize them to adjust their dividend policy accordingly. For example, when investors overvalue dividend-paying stocks, managers tend to respond with increased dividends and refrain from distributing dividends when investors prefer non-dividend companies. However, they also emphasize alternative explanations for catering behavior. For example, a dividend decrease may result from firms with higher investment opportunities retaining funds instead of reinvesting them.

Prior academics have widely discussed the concept of the market timing theory, but Baker and Wurgler (2002) were the first to label it. The theory suggests that managers tend to time equity issuances contingent on their beliefs regarding the prevailing economic conditions. Consequently, firms tend to demonstrate a market timing behavior as they tend to issue equity when they perceive the market to be overvalued and repurchase equity when it is undervalued. The connection with Jensen's (1986) hypothesis is situated at the intersection between the market conditions and capital structure determinants. Thus, a firm that issues substantial amounts of equity during prosperous times posits the risk of experiencing agency costs of free cash flow.

Moreover, Baker and Wurgler (2013) state that in behavioral finance, the market timing and the catering theories contend that a manager's investment decisions are rational responses to asset mispricing. They state two fundamental components of these theories: irrational investors must be able to cause mispricing, and managers must be "smart." The rationale is that if investors act rationally and assets are priced correctly, managers will neglect their sentiment, unable to identify and exploit mispricing opportunities created by irrational investor behavior.

2.3 The Prospect Theory & Managerial Bias

Contrary to the belief that managers are rational, behavioral finance also explores the case where managers exhibit irrational behaviors or biases themselves. Initially introduced by Kahneman and Tversky (1979), the prospect theory conceptualizes the influences that shape investment decisions in uncertain conditions. It proposes two key assumptions: Individuals tend to be loss averse and evaluate their decisions based on the potential outcomes, not just the prospective outcome, and they are more likely to be risk-averse to avoid losses than enticed by gains, known as framing. According to Kahneman and Tversky (1979), framing emphasizes that individuals' perception is affected by how the situation is framed, potentially resulting in systematically different preferences and thus investment decisions. Furthermore, reference points, such as historical investment decisions, can also guide managers' decisions as they facilitate comparing and evaluating potential outcomes. Reference points can substantially influence managers' perception of gains and losses because if a project is anticipated to yield returns above their

reference point, it is an option to pursue. Consequently, they have the potential to shape firms' investment policies.

Pertaining to the prospect theory, the concept of bounded rationality by Simon (1955) challenges the assumptions of full rationality made by conventional theories in finance. Bounded rationality encompasses the notion that when managers make decisions, their rationality is "bounded" by their cognitive limitations, restricted time frame, and information at their disposal. These challenges usually lead to applying heuristics as a mechanism for coping with complex decisions, giving rise to biases (Baker & Wurgler, 2013). For instance, a manager might be prone to choose less risky investment projects due to their risk-averse behavior.

Scharfstein and Stein (1990) developed a theoretical model to examine managerial behavior concerning investment decisions. Their model suggests that, under specific circumstances, management tends to make investment decisions based on the market sentiment and “follow the herd”, even when they possess private information that might suggest a different course of action. In this particular setting, the notion of managers neglecting their private and superior information coincides with the concept of bounded rationality. Baker and Wurgler (2013) also suggest that managers adopt “herd behavior”, albeit for a different reason, to strategically signal their optimism regarding the company’s future to investors. They assert that managers give significant weight to investors’ beliefs as it dictates the stock price today.

2.4 The Role of Cognitive Biases & Herding Behavior Theory

Transitioning from the perspective of individual investors or managers to the collective investor sentiment, Barberis, Schleifer, and Vishny (1998) were the first scholars to establish a theoretical framework for investor sentiment. Their model investigates cognitive aspects of investors’ behavior that can cause anomalies in fundamental beliefs, producing temporarily abnormal returns. For instance, they assert that stocks favored by analysts tend to yield negative returns upon announcing favorable earnings due to over-optimistic expectations resulting from such favoritism. This aligns with Griffin and Tversky’s (1992) theory, suggesting that investors generally overreact to information they deem important, such as earnings announcements, while

neglecting other vital information, particularly in response to unforeseen news. Both theories therefore present valuable insights into the effect of investors' cognitive biases that may lead to market inefficiencies.

The theoretical frameworks by Griffin and Tversky (1992) and Barberis et al.'s (1998) are associated with the herding behavior theory. According to Bikhchandani, Hirshleifer, and Welch (1992), herding behavior originates from informational cascades. They suggest that informational cascades emerge when investors disregard their private information and blindly mimic the observed actions of other investors, perceiving their actions as indicative of profitable investment opportunities, thus triggering herding behavior. However, in perfect capital markets, such behavior should be less prevalent as investors are expected to make rational decisions based on all available information. Nevertheless, Bouri, Gupta, and Roubaud (2019) argue that social media and the internet have significantly accentuated herding behavior since sharing information via these platforms has become more straightforward.

2.5 Critique & Limitations of the Theoretical Framework

A subset of scholars, including Baker and Wurgler (2002), Zhu et al. (2018), and Du and Hu (2019), emphasize the relevance of investor sentiment to corporate decisions. Their research implies that traditional theories, such as the pecking order and trade-off theories, might not sufficiently account for this aspect. Baker and Wurgler (2004) suggest that the market timing and catering theories present a more plausible explanation for the influence of investor sentiment on corporate decisions, as the pecking order and trade-off theories neglect the behavioral aspects of this process. Baker and Wurgler (2002) also suggest that managers can consistently time their decisions concerning market conditions.

Kayhan and Titman (2007) oppose the market timing theory by emphasizing that other factors like investment needs can cause significant departures from this approach. Baker, Stein, and Wurgler (2003) support Kayhan and Titman (2007) as they argue that it is unrealistic to assume that managers consistently act in the best interests of their shareholders. Baker and Wurgler (2004) provide evidence suggesting managers exhibit shareholder value-maximizing behavior.

Nonetheless, as previously mentioned, Baker and Wurgler (2004) acknowledge that other factors might affect managers' decisions, emphasizing that the catering theory is not a substitute for other theories but rather a complement.

Moreover, Kahneman (2003) claims that psychological theories in finance, such as the prospect theory, have repeatedly been criticized for its lack of a consistent alternative to the rational-agent model. However, Kahneman (2003) counteracts this by highlighting an essential aspect of the rational-agent model: Agents generally act on instincts by the visual input they perceive within a specific frame or reference at a particular moment. He also argues that this varies, both individually and collectively, thus suggesting that the prospect theory may not provide a comprehensive explanation for corporate investment decisions since it focuses on the decisions of individuals.

3. Empirical Review

3.1 Unveiling the Influence of Catering & Market Timing on Corporate Investments

Baker and Wurgler's (2004) work on catering theory initiated further research, primarily in dividend policies and earnings management, consistently supporting the premise that firms cater to investor sentiment (Rajgopal, Shivakumar & Simpson, 2007; Neves, 2017; Chazi, Theodossiou & Zantout, 2018). Extending the theory to corporate investments remains a largely unexplored topic. In the context of asset substitution, the findings by Chao et al. (2022) corroborate with previous findings, reinforcing the notion that firms strategically cater to investors through investment decisions.

Polk and Sapienza (2009) find a significant correlation between stock market valuations and corporate investment decisions. They find that this relationship is more distinct in periods of noticeably positive investor sentiment and that firms tend to cater more during these periods. However, Polk and Sapienza (2009) question the conventional views of the market timing theory: the issuance of overvalued stocks is essential to investment decisions since the proceeds are used to fund them (Baker et al., 2003). Instead, they disregard the "equity channel" in favor

of the “catering channel,” suggesting that managers tend to respond with short-term catering behavior when investors misprice stocks based on their assessment of firms’ historical investment decisions.

Kusnadi and Wei (2017) find that the market timing theory, through the equity-financing channel, can help explain the dynamics of corporate investment decisions, especially in countries with well-developed financial markets as they generally have better shareholder protection. Kusnadi and Wei (2017) also suggest that the catering theory plays a vital role in these markets due to lower costs of capital as it facilitates firms to adjust their behavior according to investors’ preferences and hence engage in catering behavior. The findings by Liao and Errico (2022) further support the concept of the catering channel. They suggest that managers, who may be shareholders, are incentivized to cater to investor sentiment through investments to influence the market valuation.

Although most prior studies support the catering and market timing theories, some researchers have yet to reach the same conclusions. Baxamusa (2011) finds that managers invest more when convinced they have profitable investment opportunities and less when uncertain about their future earnings. Consequently, this implies that managers do not cater to investors as they neglect market mispricing, thereby challenging the predictions of these theories. However, recent studies argue against Baxamusa (2011). Gu and Zhang (2022) find that managers systematically cater their investments in favor of the community’s long-term interests, thereby indicating a shift away from the traditional objective of maximizing shareholder wealth. Their study also reveals that managers may engage in catering behavior to signal valuable information to shareholders through increased investments.

Conversely, Bilel and Mondher (2021) find that firms’ catering behavior can negatively affect investments due to weakened stock performance, increasing the cost of equity, which may discourage firms from pursuing investment opportunities. Intriguingly, Chao et al. (2022) find that, despite significant uncertainty spikes, the corporate investment level did not change dramatically during the Global Financial Crisis. In other words, firms kept their long-term

investment strategies intact by not catering to investors, partially conforming with Baxamusa's (2011) and Bilel and Mondher's (2021) findings. Contrarily, Al-Thaqeb, Algharabali, and Alabdulghafour (2022) find that increases in the economic policy uncertainty index, proxied as a risk factor, positively correlate with reduced investments and market instability.

3.2 The Hidden Power of Managerial Biases on Corporate Investments

As we pivot from the catering and market timing theories, we examine the individual idiosyncrasies of managers' decisions and their influence on investment policies. Bertrand and Schoar (2003) redirect the attention to investigating the systematic behavioral differences in managers' decisions. They find significant heterogeneity across managers, especially in their personal characteristics, such as education, that appear to have a notable effect on shaping their approach to investment policy decisions. Consequently, they contend that inherent personal characteristics could explain variations in managers' behavior. However, they cannot precisely assess the effect of this phenomenon.

Expanding on this idea, Baker and Wurgler (2013) note that a bias of bounded rationality is seemingly a credible explanation for the fundamental aspects and principles behind commonly accepted standards, such as the NPV and IRR, when evaluating investment decisions. Consequently, they suggest that the discount rate of each project is tailored to its specific conditions. Conversely, Graham and Harvey (2001) argue that firms tend to deviate from these practices, suggesting that firms employ consistent discount rates across all projects, consistent with heuristics. Therefore, they suggest this causes distorted investment practices as firms may underinvest in riskier investments since they use a lower discount rate than required.

Ferris, Javakhadze, and Rajkovic (2017) explore the implications of CEO social capital on corporate decisions and policies, which they define as information, trust, and norms intrinsic to a social network. While they find that CEO social capital has a positive and significant influence on the investment policies of a corporation, as it can mitigate agency problems and encourage investments, they also demonstrate that CEO social capital positively correlates with managerial risk tolerance.

3.3 How Investor Sentiment Shapes Corporate Investments

Moving from the individual idiosyncrasies of managers to the broader market, we explore the influence of investor sentiment on corporate investment. The foundation of this area of research is established by the theoretical framework by Barberis et al. (1998) for investor sentiment. Advancing on this concept, Polk and Sapienza (2002, 2004) examine the level to which the inefficiencies of the financial markets influence managers' decisions on corporate investments. These studies provide evidence of a positive relationship between investor sentiment and the magnitude of corporate investments, indicating that higher investor sentiment leads to higher investment levels by individual firms. In contrast, Baker and Wurgler (2007) widens the scope, assessing investor sentiment using a composite measure including trading volume and dividend premiums. The findings, however, conform to those of Polk and Sapienza (2002, 2004).

Researchers concur that investor sentiment is time-varying, but there are conflicting perspectives on its duration. Some believe it has a short-term effect (Tetlock, 2007; Audrino, Sigrist & Ballinari, 2020), while others propose perseverance (Baker & Wurgler, 2007; Jiang, Lee, Martin & Zhou, 2019). However, Brown and Cliff (2005) advocate a long-term perspective as they recognize sentiment as a force for more persistent trends. This persistent effect raises concerns regarding the rationale behind managers' and investors' investment decisions as it can potentially trigger market bubbles and crashes.

Li and Grundy (2010) confirm Brown and Cliff's (2005) idea that investor sentiment shapes assets' fundamental values and further find that it also affects firms' investment levels, implying that increased investor optimism leads to increased investments. Although Arif and Lee (2014) accentuate the significance of investor optimism on corporate investment decisions, they also find that the magnitude of firms' investment positively correlates with business cycles and macroeconomic trends, reflected by GDP growth rates and analysts' long-term earnings projections. Thus, they extend the literature by showcasing how corporate investments reflect investors' expectations of upcoming economic fundamentals at both micro and macro levels.

Similar to prior studies, there is a consensus in more recent research that investor sentiment generally affects the magnitude of corporate investments (Zhu et al, 2018; Du & Hu, 2019; Li & Grundy, 2023). However, they emphasize different aspects to explain this correlation. More specifically, Zhu et al. (2018) focus on the characteristics of the top managers among Chinese firms, Du and Hu (2019) concentrate on the rationale of managers' investment decisions across companies operating in the U.S., and Li and Grundy (2023) incorporate aspects of ownership structures in the U.S.

Du and Hu (2019) provide empirical evidence that managers are susceptible to investor sentiment as they find a positive relationship between positive investor sentiment and managerial optimism. To some extent, the findings by Zhu et al. (2018) support this as they imply that younger managers have specific psychological and behavioral traits, such as being more aggressive, which makes them more inclined to respond to investor sentiment. Li and Grundy (2023) also note a positive relationship between high investor sentiment and overly optimistic managers, which could be mitigated through effective corporate governance. However, they further observe that firms with more retail investors generally invest more than comparable firms, particularly during positive sentiment.

Moreover, it is also worth noting that these studies differ in their approach to measuring investor sentiment. Du and Hu (2019) and Li and Grundy (2023) employ Baker and Wurgler's (2006) measurement, which has the limitation of being pro-cyclical. On the contrary, Zhu et al. (2018) utilize a momentum index to show investors' irrationality degree, which they note is particularly challenging to quantify. Thus, the generalizability of the results between these studies could be questioned.

3.4 Leveraging Textual Analysis to Gauge Investor Sentiment

Earlier discussions highlighted the discrepancies among researchers regarding the duration during which investor sentiment influences are noticeable. However, Brown and Cliff (2005) argue that it can have short-term effects but the long-term effects are more logical and likely. To

further this exploration, we focus on whether there is a similar pattern concerning positive and negative sentiment.

Tetlock (2007) was the initial researcher to provide empirical evidence that emphasizes the importance of social media regarding its influence on investor sentiment. Tetlock (2007) also found that weak and negative words have a more significant impact on stock market returns than positive and strong words due to negativity bias. Tetlock (2007) utilizes the General Inquirer (“GI”) to compose a straightforward measure of media pessimism derived from the coverage in the Wall Street Journal’s (“WSJ”) “Abreast of the Market” on U.S. stock market returns. Extending upon this line of research, Tetlock et al. (2008) instead utilize a quantitative measure, or textual analysis method, departing from the approach by Tetlock (2007). However, both studies used the GI to identify relevant words from the Harvard IV-4 Dictionary (“HV4”) and the Lasswell Value Dictionary (“LVD”) as a basis for their studies. Nevertheless, the authors strengthen the evidence of Tetlock (2007) that negative investor sentiment is positively correlated with abnormal returns and exerts a more substantial impact than positive investor sentiment.

Reflecting on the methodologies by Tetlock (2007) and Tetlock et al. (2008), Da et al. (2015) also utilize the GI. However, by aggregating households’ internet search volume data via Google Trends, Da et al. (2015) construct the FEARS index as a different means of measuring investor sentiment. They argue that this is a more transparent and accurate sentiment measure than previous measures because it is based on a more extensive dataset and is constructed to consider the time-varying nature of investor sentiment. Still, their findings imply that negative words exhibit the most significant impact on stock returns, guiding their decision to construct an index exclusively from terms with the strongest negative correlation with market returns.

However, some academics criticize the methodology of the studies above. Loughran and McDonald (2011) suggest that specific negative terms in the HV4 and LVD are not inherently negative in the scope of financial markets, such as “cost” or “tax”. The authors argue that some of the words that the HV4 and LVD classify as negative have neutral connotations and are used

by managers in 10-K reports, for example, to illustrate their company's operations. Nevertheless, a more recent study by Loughran and McDonald (2020) suggests that textual analysis can be an essential approach for future financial researchers. They argue that it could introduce new opportunities to measure critical economic variables that traditionally have been complex or unattainable when employing a quantitative method.

4. Literature Summary & Hypotheses Development

The empirical review presented in [Appendix 1](#) underscores a consensus in prior research that investor sentiment positively correlates with the magnitude of corporate investments (Baker & Wurgler, 2007; Polk & Sapienza, 2009; Li & Grundy, 2010, 2023; Zhu et al., 2018; Du & Hu, 2019). The conformity of these findings suggests that retail investor sentiment can serve as a credible predictor of firms' investment behaviors. However, while traditional economic theories underscore the importance of concepts such as NPV and IRR to understanding the rationale behind corporate investment decisions, these findings hints at other determinants in play, including the influence of psychological or behavioral elements.

Upon reviewing the nuances of corporate investment, it is evident that the market timing theory, catering theory, and behavioral aspects are relevant in connecting investor sentiment to firms' investment decisions. Despite this, alternative perspectives exist. Baxamusa (2011) recognizes managers' perception of the prevailing market conditions as the primary driver of firms' investment decisions instead of sentiment. Interestingly, Chao et al. (2022) found that firms did not abandon their investment strategies even in prevailing uncertainty during the 2007 financial crisis, indicating a potential reluctance to adhere to prevailing investor sentiment. Notably, this contradicts the prospect theory and managers' bounded rationality while aligning with Baxamusa (2011). Thus, if managers are risk-takers when facing losses and being influenced by their cognitive biases, low retail investor sentiment should lead to decreased investment activities. Furthermore, although Chao et al. (2022) uphold the market timing theory, they challenge the catering theory's predictions. Yet, this aligns with Baker and Wurgler's (2004) alternative explanations, and the empirical evidence supports these theories (Zhu et al., 2018; Du & Hu, 2019; Li & Grundy, 2023), consistently indicating that investor sentiment positively correlates

with corporate investments. Nevertheless, the conclusions by Loughran and McDonald from 2020 question whether there is a definite positive correlation between investor sentiment and corporate investments, as textual analysis may provide new perspectives, previously unattainable. This forms the basis for our first hypothesis:

H1: *Retail investor sentiment, proxied by the FEARS index, is predictive of changes in firms' investment activities*

Considering H1, Chao et al. (2022) found that during a global crisis, the fundamental elements of firms' investment decisions precede the psychological aspects of managers' behavior, suggesting a relatively low fluctuation in investment levels during crisis periods. Similarly, our study period encompasses the pandemic, which drastically altered the landscape of financial markets (Al-Thaqeb et al., 2022). This shift in investor sentiment and behavior, reflected by changes in the FEARS index, raises questions about firms' investment decisions during crises. However, as Chao et al. (2022) noted, fundamental factors related to investment opportunities tend to overshadow shifts in investor sentiment, suggesting a possible decrease in the influence of sentiment on firms' investment activities. Nevertheless, it is essential to remember that the concepts of reference points and framing pertinent to the prospect theory could lead to systematically different outcomes, which begs the question of whether this pattern may also be counteracted during this period. Consequently, this leads us to the formulation of our second hypothesis:

H2: *The pandemic has enhanced the predictive efficacy of the FEARS index, reflecting an increase in retail investors' influence*

While the previous studies support H1, they rely only on various market-based measures of investor sentiment. These measures might yield skewed outcomes as they are subject to influences beyond the sentiment itself and often overlook an increasingly important share of investors, namely retail investors (Da et al., 2015). This study's main emphasis is on the FEARS index, a measure grounded in Google search data, which offers a more accurate and transparent

proxy for retail investor sentiment. Should a significant correlation between the FEARS index and firm investment exist, it could imply that the influence of retail investors exceeds what has been traditionally acknowledged by conventional theories in finance. Consequently, as Bouri et al. (2019) find that the emergence of the internet and social media has led to more significant market inefficiencies, specifically information asymmetries, we anticipate the FEARS index to exhibit a stronger correlation with corporate investment than prior measures. All else equal, more significant information asymmetries should result in more pronounced variations in investor sentiment as they correlate with assets' mispricings which, in turn, should lead to greater variations in firms' investment levels. However, using a search-based measure like the FEARS index in the context of firms' investment activities is largely unexplored. Thus, whether the FEARS index exhibits a similar pattern as market-based measures pave the way for our third hypothesis:

H3: *The strength of the relationship between the FEARS index and corporate investments differs from that observed with market-based measures of investor sentiment, such as the VIX.*

5. Data & Sample Description

5.1 Sample Description

Our empirical analysis utilizes a sample of publicly listed firms on the Nasdaq Composite ("CCMP") to examine the relationship between retail investor sentiment, measured by the FEARS index, and corporate investment policies. The CCMP is a broad-based capitalization-weighted index consisting of small and medium enterprises and large enterprises across all three Nasdaq tiers: Global Select, Global Market, and Capital Market (Bloomberg Finance L.P.) It has a notable concentration of firms operating in Technology, Healthcare and Biotechnology, and Consumer Services and Retail sectors while still encompassing a diverse range of companies across various industries ([Appendix 2](#)).

The financial statement data is collected from Bloomberg on a quarterly frequency from 2013 to 2022, i.e., from Q1-2013 to Q4-2022. Due to our study's design, our empirical analysis period

ranges from Q1-2014 to Q4-2021 to account for future effects. Quarterly data more effectively align with investor sentiment dynamics, which can experience substantial short-term fluctuations, making annual data less appealing (Brown & Cliff, 2005). The time period is selected to encompass different economic cycle phases to account for fluctuations in economic activity.

Firms with incomplete data were excluded to ensure analytical consistency across observations and minimize biases from imputing missing values. Information regarding firm age was collected manually through Google Finance. Following Lee and Loughran (1998) and Chao et al. (2022), we adjust our final sample by excluding American Depositary Receipts and Real Estate Investment Trusts. For dual-class shares, we select one class representing each company in the sample. Finance, Insurance, and Real Estate (SIC codes 6000-6999) and Public Utilities (SIC codes 4910-4949) are excluded from the sample because their financial structures and regulatory environment differ from most industries. To address the presence of outliers, we winsorize the top and bottom 1% of all accounting-related variables. The final sample consists of a strongly balanced panel of 23,712 quarterly observations on 741 individual firms.

We select the U.S. as our country of study for several compelling reasons. Firstly, the U.S. offers a homogenous regulatory environment under the unified framework of the Securities and Exchange Commission ("SEC"), as acknowledged by Lobo (2017). The high degree of uniformity in the financial regulatory framework makes the U.S. an appropriate setting for our research. In contrast, Europe's regulatory landscape is more fragmented due to multiple regulatory bodies across its member countries. The uniform reporting practices under the SEC result in standardized reporting across all firms, ensuring a consistent and reliable data sample for our empirical research (SEC, 2023). This facilitates the extraction of insights that may not be as easily discernible in smaller or less developed markets.

The U.S. has a distinctive investment culture marked by comparatively high levels of individual participation in the stock market relative to other nations (Guiso, Sapienza, & Zingales, 2008; Deloitte, 2021). Since we hypothesize that the power and influence of retail investors in financial

markets are growing, the U.S. is an ideal candidate to assess whether firms cater to retail investor sentiment through their investment policies.

5.2 Variable definition

In the following section, i indicates the study sample of firms, and t denotes the period measured in quarters unless stated otherwise. [Appendix 3](#) offers a summary of all variables. The details and construction of our main regressor variable is presented in section 6. To facilitate the interpretation of our results and to provide economic meaning in section 7, we will adjust our interpretation to reflect the dependent variable in percentage terms.

5.2.1 Dependent Variable

5.2.1.1. *Investment activities* (“*IA*”) is used as the main dependent variable to examine the relationship between investor sentiment and corporate investment policies. Following the study by Chao et al. (2022), the magnitude of investment activities is defined as follows:

$$IA_{i,t-1} = \frac{(I_{i,t} + I_{i,t+1} + I_{i,t+2})/3}{Total\ assets_{i,t-1}}$$

Investments for the sample firms ($I_{i,t}$) include the amount spent on purchases of tangible and intangible fixed assets, specifically capital expenditures and R&D spending. Considering the time lag between investment decisions and their execution, investment activities equals the average investments across the subsequent three quarters divided by total assets¹ to accurately capture the impact of these decisions and accommodate potential delays in their implementation. Normalizing the measure by total assets allows for meaningful comparisons across firms of different sizes and industries. The variable *IA* is expected to provide insight into how firms allocate resources in response to investor sentiment.

¹ Total assets is defined as the book value of total assets.

5.2.2. Independent variables

5.2.2.1 The *Financial and Economic Attitudes Revealed by Search (“FEARS”)* index is the main explanatory variable in the regression model, and it is intended to capture the effect of changes in retail investor sentiment. The FEARS index reveals the collective economic sentiment through the analysis of Google search queries of millions of households in the U.S. It approximates the aggregate sentiment of retail investors by tracking the volume and frequency of predetermined search terms related to financial markets and economic uncertainty. Polk and Sapienza (2009) found that high investor sentiment positively correlates with increases in corporate investments, hence the relationship between investor sentiment and investment activities is expected to be positive. The FEARS index is calculated as the change in adjusted search volume for the thirty search terms with the strongest relationship to market returns, ranked by their absolute t-statistic values:

$$FEARS_t = \sum_{i=1}^{30} R^i(\Delta ASVI_{j,t})$$

where $R^i(\Delta ASVI_{j,t})$ is the $\Delta ASVI_{j,t}$ for the search term with a t-statistic rank of i and j denotes search terms. An increase in the variable value indicates positive sentiment, and a decrease indicates negative sentiment.

5.2.2.2. *Free cash flow (“FCF”)* accounts for differences in internal resources available for investments. Fazzari, Hubbard, and Petersen (1987) and Cleary (1999) find that firms with higher financial constraints exhibit greater sensitivity of investment to cash flow, and Jensen’s (1986) hypothesis suggests that excessive free cash flow may lead to overinvestment. The relationship between FCF and IA is thus expected to be positive. Consistent with Chao et al. (2022), FCF is calculated as the cash flow from operations divided by total assets at the beginning of the year. Normalizing the measure allows for meaningful comparisons across firms of different sizes and industries. However, to smooth out short-term fluctuations in cash flows and better capture the underlying trends, the calculation includes the trailing twelve months

(“TTM”) cash flow from operating activities divided by total assets the quarter prior to the TTM period:

$$FCF_{i,t} = \frac{TTM \text{ Cash from operations}_t}{Total \text{ Assets}_{i,t-5}}$$

5.2.2.3. *Debt ratio (“DR”)* accounts for the influence of firms’ leverage on their investment decisions. Empirical evidence shows that high debt ratios can restrain managers from engaging in non-profit maximizing behavior, thereby reducing overinvestment (Myers, 1977; Jensen, 1986; Aivazian, Ge, Qiu, 2005). Lang, Ofek, and Stulz (1996) and Rajan and Zingales (1995) found that leverage negatively affects investment opportunities, suggesting that higher debt levels may limit firms’ investment ability. Furthermore, Aivazian et al. (2005) found that this negative effect is significantly stronger for firms with low growth opportunities, supporting the notion that leverage functions as a disciplinary mechanism. Thus, DR is expected to have a negative relationship with investments. The DR is measured as total debt relative to total assets:

$$DR_{i,t} = \frac{Total \text{ Debt}_{i,t}}{Total \text{ Assets}_{i,t}}$$

5.2.2.4. *Firm size (“Size”)* accounts for different investment behaviors between firms of different sizes. Larger firms may have better access to financing, more diversified operations and economies of scale, whereas smaller firms may face more financial constraints and greater uncertainty, indicating more cautious investment strategies. Bond and Meghir (1994) find that larger firms exhibit different investment patterns than smaller firms. Cleary (1999) finds a positive relationship between size and investment, indicating that larger firms tend to invest more, whereas Lang et al. (1996) suggest that larger firms have lower growth opportunities and are more likely to decrease their investments. However, as consensus points toward a negative relationship, the coefficient of size is expected to be negative. In the summary statistics in section 5.3, size is unscaled to provide economic meaning, but defined as the natural logarithm of total assets in the regression model to improve its normal distribution:

$$Size_{i,t} = \ln(Total\ Assets_{i,t})$$

5.2.2.5. *Return on Assets (“ROA”)* accounts for the influence of firms’ financial performance on their investment decisions. ROA functions as a proxy for profitability, measuring how efficiently a company utilizes its assets to generate earnings. Fazzari et al. (1988) and Cleary (1999) provide evidence of a positive relationship between ROA and corporate investment decisions, indicating that more profitable firms tend to invest more. Consequently, the relationship between ROA and investing activities is expected to be positive. ROA is measured as the net income divided by total assets.

$$ROA_{i,t} = \frac{Net\ Income_{i,t}}{Total\ Assets_{i,t}}$$

5.2.2.6. *Market-to-book ratio (“MB”)* is used as a proxy to account for differences in growth potential and investment opportunities across firms. Penman (1996) used the MB to indicate firms’ growth opportunities, as it captures the market’s expectations of future growth relative to the book value of equity.. The relationship between MB and investment is expected to be positive. The MB is calculated by dividing the market value of equity by the book value of equity:

$$MB_{i,t} = \frac{Market\ Value\ of\ Equity_{i,t}}{Book\ Value\ of\ Equity_{i,t}}$$

5.2.2.7. *Firm age (“Age”)* accounts for the potential influence of a company’s lifecycle stage on investment decisions, as younger and older firms typically have different investment patterns, financial constraints and growth opportunities. Gomes (2001) finds that young firms tend to invest more aggressively due to their higher growth opportunities and are more sensitive to financial constraints than older firms, suggesting that firm age plays a role in determining investment behavior. The relationship between Age and investment is expected to be negative. Age is defined as the number of years since its founding as of the start of the year.

5.2.2.8. *Board size* (“*BoardSize*”) accounts for the potential influence of a company’s board structure on investment decisions. Board size can affect the quality of decision-making and the ability of the board to monitor management effectively. Larger boards are associated with decreased investment inefficiency, suggesting that larger boards can positively affect corporate investments (Chen & Jaggi, 2000). Consistent with Chao et al. (2022), *BoardSize* is defined as the total count of directors serving on the board.

5.2.2.9. *Board Independent* (“*Ind.*”) accounts for differences in corporate governance quality and potential conflicts of interest between the board and management. Independent directors are crucial in monitoring management and ensuring their decisions align with stakeholder interests. Boone, Field, Karpoff, and Raheja (2007) and Baysinger and Butler (2019) find that more independent boards are less likely to make suboptimal investments, suggesting that board independence can positively influence investment decisions. Thus, a higher degree of board independence is associated with better investment decisions as independent directors can more effectively monitor and challenge management, providing better shareholder protection. Consistent with Chao et al. (2022), *Ind* is measured as the proportion of independent directors on the board.

5.2.2.10. *Chair Duality* (“*Dual*”) is a dummy variable that accounts for governance structures in which the CEO, or equivalent, also serves as the Chair of the board. The chair duality structure may enhance managerial discretion, enabling the board’s chairman to control the information available to other board members, potentially hindering effective monitoring and oversight (Jensen, 1993). Consequently, the presence of chair duality should influence the magnitude of investment. The dummy variable is assigned a value of one if the CEO is also the chair of the board and zero otherwise.

5.3 Summary Statistics

[Table 1](#) reports summary statistics for the variables utilized in the empirical analysis. The dependent variable, investment activities, ranges from 0 to 6.1%, implying that some firms had no capital expenditures in some quarters. In contrast, other firms have invested an equivalent of

6.1% of their total asset base. The sample firms, on average, invest 1% of their total asset base over the next three quarters, and the standard deviation of 0.01 further suggests that this is the case for most firms in the sample. The average sample firm is cash flow positive, as seen by the mean of 0.01. The distribution is, however, negatively skewed, suggesting that most firms have relatively higher free cash flow generation. The wide range of -1.4 to 0.43 highlights the variation in cash-generating abilities across the sample firms. This can be explained by many young Technology and Healthcare firms listed on the CCMP, which tend to be unprofitable in their early growth stages. The average firm has total assets equivalent to \$3563.69m. The right-skewed distribution suggests that most firms have relatively lower total assets. However, the presence of a long right tail reveals that there are firms with substantially larger asset bases, likely belonging to capital-intensive industries or holding dominant market positions. This aligns with intuition as several mega caps are listed on the CCMP. Total assets range from \$7.23m to \$77,093m, meaning that the largest firm is more than 10,000 times larger than the smallest firm. To account for these large-size differentials, complimentary summary statistics by size groups are presented in [Appendix 4](#).

Leverage across the sample firms ranges from 0 to 1.25, indicating that some firms are debt-free. The max-observation is due to negative equity in the company, indicating that it is overleveraged. The average firm has leverage equal to 22% of its assets base. The distribution is skewed to the right, suggesting that most firms have relatively lower debt ratios, indicative of conservative capital structures and lower reliance on debt financing. The average sample firm generates a net return of -6% on its asset base. The left-skewed distribution suggests that most firms generate relatively higher returns on their asset bases, indicating higher operational efficiency and profitable use of assets. However, the presence of firms with substantially lower, negative returns pulls down the sample mean. This is seen in the range of observations that varies from -149% to +37%. However, this is likely explained by the young and unprofitable cluster of firms listed on the CCMP, as these firms often have scalable business models associated with high fixed costs. This is further observed as the average sample firm is 39.78 years old. The right-skewed distribution suggests that most firms are relatively younger and potentially in growth or expansion stages. The average sample firm has a market-to-book ratio

equivalent to 4.09. The right-skewed distribution suggests that most firms have market valuations closely aligned with their book equity values. However, as seen in the range from -22.65 to 46.96, there is a significant variation in market valuations and growth prospects across the sample firms, further emphasized by the relatively high standard deviation.

Most firms have similar corporate governance structures as the average and median board size converges around eight directors. This could indicate a consensus on optimal board size for effective corporate governance. The board independence average equals 82% with a standard deviation of 9%, meaning most observations are clustered somewhat close to the mean. The distribution is slightly skewed to the left, suggesting that most firms have a higher proportion of non-executive directors on the board. This indicates that firms prioritize board independence, typically associated with better corporate governance practices. The limited variance across firms suggests that most firms adhere to regulatory requirements or general best practices. Lastly, CEO duality has a mean of 0.34, indicating that the CEO is also the board chairman in one-third of the sample firms.

6. Methodology

6.1 The FEARS Index Proxy

The FEARS index is considered an effective proxy of retail investor sentiment for several reasons. First, it is a representative representation of retail investor sentiment as many individuals use Google to gather financial information, providing a less filtered view of sentiment as users typically search for exactly what concerns them. Furthermore, the quantifiable nature of the FEARS index allows for consistent application over different periods and market conditions. However, the proxy captures the sentiment of all search engine users, introducing potential noise, and using a proxy introduces potential measurement errors when there is a disparity between the actual variable of interest and the proxy. Nonetheless, based on household internet search data, it is likely to reflect the sentiment of retail investors specifically.

6.1.1 Google Trends

Google Trends is an online tool made available by Google that enables users to examine and visualize the search volume of specific terms across regions and periods. Consequently, Google Trends allows us to direct our data collection to only include search volume data in the U.S. from 2013 to 2021. It normalizes search data to enable comparability over time and across terms, producing a Search Volume Index (“SVI”) ranging from 0 to 100, with 0 indicating the lowest search volume and 100 the highest search volume (Google Trends, 2023). To maintain consistency with the firm-level data, we convert the monthly SVI data into quarterly frequency by averaging the SVI of each quarter’s three months.

6.1.2 Collecting & Adjusting the Google Trends Data

Following the methodology employed by Da et al. (2015), the first step is to construct a list of search queries that reveal sentiment toward economic conditions using the HV4 and LVD dictionaries². These dictionaries categorize words into groups such as “Positive”, “Pstv”, “Negative”, and “Ngtv”. Because we are interested in investor sentiment, we select the “Economic” words classified as “Positive” and “Negative” while excluding neutral words. This yields a primitive word list consisting of 138 search terms such as “rich”, “gold”, and “crisis”. Subsequently, to discern how households may use these terms in Google searches, we input each primitive word into Google Trends to identify its top ten related search queries. Thus, our primitive word list generates 1,380 related search terms, which become 1,350 terms after eliminating duplicates. Next, we remove terms with insufficient data, i.e., if the search term displays an SVI of 0 in more than ten quarters. Finally, terms without economic substance are removed. For instance, searching for “bankrupt” yields related searches such as “bankruptcy” and “bald and bankrupt”. Thus, we keep the former (related to bankruptcies) and remove the latter (a youtube-channel), which ultimately yields a list of 197 search terms.

² Our study encountered challenges due to the discontinuation of the GI. The index has therefore been manually constructed. We sourced the HV4 and LVD wordlist from the Wayback Machine, which appears outdated as it does not include terms related to Covid. Updating the list independently would compromise the methodology’s credibility and introduce potential bias, as we do not know the selection criteria used in constructing the wordlist.

Next, we download the SVI for each of the 197 search terms over our sample period from October 2012 to December 2022 from Google Trends. Consistent with Da et al. (2015), we define the quarterly change in each of the search terms as:

$$\Delta SVI_{j,t} = \ln(SVI_{j,t}) - \ln(SVI_{j,t-1})$$

where j indicates the search terms, and t denotes the period measured in quarters. To provide an example, [Figures 1](#) and [2](#) illustrate the quarterly log changes in the search terms “Silver” and “Charitable donation” during our sample period. Seasonality becomes evident, as the SVI consistently spikes in Q1 every year and subsequently declines throughout the year. Further, there are substantial differences in variance across the search terms, indicative of heteroskedasticity. Finally, the SVI change for “Charitable donation” indicates the presence of some extreme values. These observations warrant further data treatment of the change in search term volumes.

We adjust the raw data by winsorizing it at the 5% level to mitigate concerns about outliers. To address the seasonality issue in $\Delta SVI_{j,t}$, we regress $\Delta SVI_{j,t}$ on quarterly dummies and retain the residual. Finally, to address heteroskedasticity and enable comparability across each time series, we follow the method of Baker and Wurgler (2006) and standardize each time series by scaling it relative to its standard deviation. This yields an adjusted quarterly change in search volume, $\Delta ASVI_{j,t}$, for each of our 197 search terms.

6.1.3 Constructing the Index

The concluding step in constructing the FEARS index is to identify the search terms with the most significant impact on market returns in a given quarter. We mirror the method employed by Da et al. (2015) and run expanding backward rolling regressions of $\Delta ASVI_{j,t}$ on market returns every second quarter (June and December) to establish the historical relationships between our search terms and the CCMP returns. Da et al. (2015) find that search terms that strongly correlate

with the market generally exhibit a negative relationship³. Consequently, they only utilized the terms with the largest negative t-statistic on $\Delta ASVI_{j,t}$ to construct their FEARS index.

[Table 2](#) shows the descriptive statistics of the t-statistics from the regression of $\Delta ASVI_{j,t}$ on market returns for all our search terms. Unlike the findings of Da et al. (2015), the t-statistics on $\Delta ASVI_{j,t}$ in our sample exhibit a more normal distribution ([Figure 3](#)) with a mean of -0.023 and a minor rightward skew with significant outliers in both tails of the distribution, ranging from -23.3 to 23.724 ([Figure 4](#)). Since our distribution of t-statistics fundamentally differs from Da et al. (2015), we choose to rank the t-statistics by their absolute values. Hence, we include search terms that exhibit both positive and negative relationships with market returns and construct an index based on the terms with the strongest relationship, independent of the sign. A caveat in our implementation is the relatively short time series from which the relationship is estimated. However, the regression coefficient from the search terms and market returns is merely used as qualitative guidance for selecting the most critical search terms. Ultimately, the FEARS index does not utilize the regression coefficients as direct inputs but is based on the quarterly change in the adjusted search term volumes.

Our methodology differs from that of Da et al. (2015) for two reasons. First, instead of daily frequencies, we use quarterly frequencies to align the FEARS index with the time frames of our firm-level data. Second, while we acknowledge the efficacy of negative terms in capturing immediate shifts in sentiment, we argue that both negative and positive terms play essential roles in shaping investor sentiment over the business cycle. Adverse events may trigger sharp market reactions due to negativity bias (Tetlock, 2007), but positive developments, such as robust economic growth, can also significantly sway investor sentiment. Thus, we posit that sentiment is not a one-sided coin but a dynamic interplay of positive and negative factors. In light of this, we have modified the FEARS index to include positive terms, striving for a more balanced and accurate representation of retail investor sentiment over time. Thus, by ranking t-statistics by their absolute values on $\Delta ASVI_{j,t}$, our FEARS index is defined as:

³ In their total sample from 2004 to 2011, they found zero search terms with a t-statistic on delta ASVI greater than 2.5 but identified fourteen search terms with a t-statistic below -2.5.

$$FEARS_t = \sum_{i=1}^{30} R^i(\Delta ASVI_{j,t})$$

where $R^i(\Delta ASVI_{j,t})$ is the $\Delta ASVI_{j,t}$ for the search term j with a t-statistic rank of i , ranging from the largest ($i=1$) to the smallest ($i=197$) absolute values within each quarter. We then select the thirty terms in quarter t with the strongest relationship with market returns to form the FEARS index for that period. FEARS in quarter t is thus the average $\Delta ASVI_{j,t}$ of these thirty terms in quarter t . The initial window is set to 4 quarters to further enhance the statistical power, which is why our FEARS index starts in Q1 2014. To illustrate, [Appendix 5](#) displays the top thirty terms for Q1 2014 and [Figure 5](#) illustrates the FEARS index over the selected sample period.

6.2 Univariate analysis

We start our empirical analysis by segmenting the data sample into two subsamples: the pre-Covid and the Covid periods. Utilizing univariate tests, we examine the differences in the mean values of our dependent variable, investment activities, and the main explanatory variable, the FEARS index, between these two periods. This analysis provides a preliminary indication of how the onset of the pandemic may have impacted the variables under study.

6.3 Pre-regression Diagnostics

Our study employs a multiple linear regression model to investigate the relationship between retail investor sentiment and corporate investment policies. The Ordinary Least Squares (“OLS”) method is used to estimate the model due to its interpretability and widespread acceptance in economic research. [Table 3](#) displays a correlation table for the independent variables in the model. However, some correlation coefficients are remarkably high, FCF-ROA of 0.850 in particular. Although the assumption of no perfect collinearity is not violated, these coefficients may suggest the presence of multicollinearity in the model, which needs to be investigated further. The variance inflation factor (“VIF”) shows a value of 3.89 for ROA ([Table 4](#)). A common rule of thumb is that a VIF greater than five might indicate problematic

multicollinearity. The high correlation coefficient of FCF-ROA may thus not be a problem, but we proceed cautiously, noting the potential issues in the results. Further, as the study employs time-series data, observations will be correlated over time, and serial correlation will be present.

The homoskedasticity of the residuals was assessed using White's test, which yielded a statistically significant p-value of 0.000 ([Table 5](#)). Thus, the null hypothesis of constant variance can be rejected at the 1% significance level in favor of the alternative hypothesis of heteroskedasticity, which may lead to inefficient OLS estimates of the standard errors of the coefficients. To adjust for heteroskedasticity, the hypothesis testing utilizes robust standard errors and standard errors clustered by firm, where the latter accounts for potential serial correlation over time for a given firm.

To determine whether to use random effects ("RE") or fixed effects ("FE") estimators, we utilized the Hausman test. The test yielded a statistically significant p-value of 0.000 ([Table 6](#)). Thus, the null hypothesis can be rejected at the 1% significance level in favor of the alternative hypothesis that there is a systematic difference in the coefficients between the RE and FE models. This suggests potential endogeneity in the error terms, implying that the FE model is more appropriate. The FE model allows us to control for all time-invariant characteristics that are unique to each entity and would otherwise bias our results. These characteristics, such as industry, could be observed or unobserved, such as time-invariant management culture. Thus, the FE model effectively controls for firm-specific omitted variables that don't vary through time, limiting the scope of the endogeneity problem (Denis, 2015). However, firm-specific time-varying variables may still be a source of endogeneity in our model.

6.4 Model specification

Based on the conclusions from the pre-regression diagnostics, the relationship between retail investor sentiment and corporate investment policies is examined with fixed effects panel regressions. Consistent with economic research best practices, heteroskedasticity is accounted for using clustered standard errors as it adjusts for the remaining within-firm error correlations over time. Following the studies of Polk and Sapienza (2009) and Chao et al. (2022), our multivariate

analysis model includes free cash flow, leverage, size, profitability, market-to-book, age and corporate governance specifications. Hence, we examine the relationship between retail investor sentiment and corporate investment policies from 2014 to 2021 using the following multiple linear regression specification:

$$IA_{i,t} = \beta_i + \beta_1 FEARS_t + \beta_2 FCF_{i,t} + \beta_3 DR_{i,t} + \beta_4 Size_{i,t} + \beta_5 ROA_{i,t} + \beta_6 MB_{i,t} + \beta_7 Age_{i,t} + \beta_8 BoardSize_{i,t} + \beta_9 Ind_{i,t} + \beta_{10} Dual_{i,t} + \varepsilon_{i,t}$$

where the variables are defined as in section 3.2. The coefficient β_i is the firm-fixed effect, which controls for all time-invariant characteristics of the firms in the sample⁴. For illustrative purposes, [Table 7](#) shows the model under the specifications discussed in the pre-regression diagnostics, including regular standard errors and fixed effects. As expected, the coefficients remain unchanged in the (1-3) model specifications. Model 4, including fixed effects and clustered standard errors, is the specification that will be used in the hypothesis testing⁵.

6.5 Subsample Analysis

As visible in [Figure 5](#), the initial quarters of 2018 (1.00) and 2021 (-1.18) present relatively extreme realizations of the FEARS index. Thus, to further deepen our analysis related to our first hypothesis, we focus on the specific circumstances during these periods to better understand their distinct impact on investment activities. We do this by introducing quarter dummies for these quarters in our main regression:

$$IA_{i,t} = \beta_i + \beta_1 FEARS_t + \beta_2 FCF_{i,t} + \beta_3 DR_{i,t} + \beta_4 Size_{i,t} + \beta_5 ROA_{i,t} + \beta_6 MB_{i,t} + \beta_7 Age_{i,t} + \beta_8 BoardSize_{i,t} + \beta_9 Ind_{i,t} + \beta_{10} Dual_{i,t} + QuarterDummy_t + \varepsilon_{i,t}$$

⁴ Such time-invariant firm characteristics are for example the industry in which a firm operates, since firms do not change industry throughout the sample period.

⁵ We do not control for year effects as our main explanatory variable exclusively varies over time, which would absorb its variation.

To test our second hypothesis, we introduce a period dummy variable that differentiates between the time-frame before and after the onset of the pandemic.

7. Empirical Findings & Analysis

7.1 Univariate Analysis

The empirical analysis begins with univariate tests. [Table 8](#) presents the differences in means for our dependent variable, investment activities, and our main explanatory variable, the FEARS index, before and after the onset of the pandemic. The pre-Covid period is defined as the time-frame before the first quarter of 2020, accounting for 75% of the sample observations, while the Covid period commences in the first quarter of 2020, accounting for the remaining 25% of the sample. Investment activities display a p-value of 0.000. Thus, we reject the null hypothesis and conclude that there is a statistically significant difference at the 1% level between the means of investment activities in the two periods. The mean value of investment activities changes slightly between the periods, from 0.9% to 0.8%, indicating that the sample firms, on average, decreased their investment activities during the pandemic. The FEARS index displays a p-value of 0.000. Thus, we reject the null hypothesis and conclude that there is a statistically significant difference at the 1% level between the means of the FEARS index before and during the pandemic. The mean value of the FEARS index during the Covid-period is significantly lower than its pre-Covid period mean value, changing from -0.002 to -0.156, indicating a considerably more negative retail investor sentiment during the Covid-period. Thus, as the retail investor sentiment significantly worsened during the pandemic, companies, on average, slightly decreased their investment activities during the same period. Amidst the pandemic-induced uncertainty, firms likely mimicked the actions of their peers, contributing to an overall decrease in corporate investment levels. Interestingly, this preliminary indication of the impact of the pandemic on our dependent variable goes against Chao et al. 's (2022) findings that firms' investment levels remain steady during crises.

7.2 Multivariate Analysis

7.2.1 Hypothesis 1

[Table 9](#) reports the estimation results of our primary regression Model 4, utilizing firm fixed effects and clustered standard errors. Our main explanatory variable, FEARS, exhibits a negative coefficient (-0.000), statistically indistinguishable from zero, and a p-value of 0.985. Thus, the model fails to establish a statistically significant relationship with investment activities over the full sample period, leading us to reject our first hypothesis. The economic significance would be of little consequence irrespective of a statistically significant relationship, as variations in FEARS appear to exhibit a substantially small influence on firms' investment activities.

Intriguingly, as the model fails to establish a statistically significant relationship, we note two outliers from the long-term trend in the FEARS index ([Figure 5](#)). To add nuance to our results, we examine the sensitivity of our findings to these outliers by conducting a subsample analysis. We modify the sample period by excluding the first quarters of 2018 and 2021 separately, and then simultaneously, to examine whether time-bound events are influencing our data. For instance, a specific quarter may have been subject to unusual economic occurrences or market volatility that do not accurately represent the typical market conditions during the sample period, contributing to a potential temporary discrepancy between the FEARS index and firms' investment activities.

The Tax Cuts and Jobs Act was enacted on the first of January 2018, which reduced the corporate tax rate from 35% to 21%, leading to increased corporate investments (Gale & Haldeman, 2021). In light of the emerging trade tensions between the U.S. and China, corroborating Jensen's (1986) hypothesis predictions, we posit that expectations of augmented corporate investments triggered a sharp positive shift in investor sentiment, supporting Griffin and Tversky's (1992) notion that investors typically overreact to unforeseen news. Due to the ongoing economic fallout from the pandemic, recession worries emerged in the first quarter of 2021, contributing to the short-term decline in retail investor sentiment.

[Table 9](#) reports the estimation results from Models 5, 6, and 7. In Model 5, the first quarter of 2018 is excluded. The FEARS index's estimated coefficient (-0.028) is negative and statistically significant at the 1% level. While the model provides evidence of a robust relationship between firms' investment activities and retail investor sentiment, our first hypothesis is rejected as we predicted a positive relationship. The negative coefficient value indicates an inverse relationship. Explicitly, a one unit increase in the FEARS index is, on average, associated with a 0.028-unit decrease in firms' investment activities, all else equal. Proceeding to Model 6, where the first quarter of 2021 is excluded, the FEARS index's estimated coefficient (0.016) is positive and statistically significant at the 10% level. Although this implies a weaker correlation than in Model 5, the observed relationship changes direction. It conforms with our first hypothesis that the FEARS index is predictive of investment activities, suggesting we fail to reject the hypothesis at the 10% level. A one unit increase in the FEARS index is associated with a 0.016 unit increase in firms' investment levels, all else equal.

While Models 5 and 6 pique our curiosity, we lack a theoretical justification to exclude only one of the anomalous quarters in the FEARS index. To avoid selectively choosing which data to exclude to achieve a statistically significant result, the most transparent approach is to pivot to Model 7, which excludes both the first quarters of 2018 and 2021. Nevertheless, the model fails to establish a statistically significant relationship over the adjusted sample period, leading us to yet again reject hypothesis 1. However, the estimation results have improved significantly from Model 4 to Model 7, suggesting high sensitivity to the exclusion of certain quarters. In conclusion, as our sample period encompasses two fundamentally different periods in terms of retail investor sentiment before and after the pandemic's onset, the model cannot establish a statistically significant relationship between retail investor sentiment and firms' investment levels.

7.2.2 Hypothesis 2

Our analysis progresses by revisiting the univariate tests, which showed statistically significant differences in the means of investment activities and the FEARS index before and after the

pandemic onset. Thus, we separate our analysis into pre-Covid and Covid and present the corresponding estimation results from Models 8 and 9 in [Table 10](#).

The estimation of Model 8, representing pre-Covid, yields a positive and statistically significant coefficient for FEARS (0.027) at the 1% level, suggesting a robust relationship between retail investor sentiment and investment activities. A one unit increase in FEARS is associated with a 0.027-unit increase in investment activities. Conversely, the estimation of Model 9 also yields a positive and statistically significant coefficient at the 1% level, albeit demonstrating a considerably larger magnitude of the coefficient (0.052), indicating significantly increased responsiveness to investor sentiment during the pandemic. A one unit increase in the FEARS index is associated with a 0.052-unit increase in investment activities. Economically, this indicates that the influence of retail investor sentiment has increased after the onset of the pandemic. Thus, we find support for our second hypothesis. However, since there are very few quarters in the second part of the sample, these findings could represent weak evidence and should be interpreted with caution.

In this setting, our results are consistent with earlier research (Polk & Sapienza, 2009; Baker & Wurgler, 2007; Zhu et al., 2018; Du & Hu, 2019; Li & Grundy, 2023), confirming that investor sentiment positively correlates with firms' investment activities. However, our study delves deeper, emphasizing how this relationship shifts under the pandemic, a period in which a significant increase in retail investor activity was observed, partially attributable to the transition to remote work as it allowed them to allocate more time for involvement in the stock market (WSJ, 2020; Deloitte, 2021). Theoretically, this shift should intensify the significance of the FEARS index, especially when this enhancement reflects a recent trend.

The pandemic required adaptable and innovative risk management strategies since it caused an unprecedented situation that may suggest that managers relied more on their intuition rather than heuristics, which became abruptly irrelevant when the pandemic onset. Furthermore, this may have been caused by a greater emphasis on flexibility and adaptability, given the changing nature of the crisis. The nature of the pandemic may also have increased the influence of psychological

biases and cognitive errors, underscoring the essential role behavioral aspects play in corporate decision-making. As the pandemic has shown, traditional economic models may not wholly account for decision-making processes under severe uncertainty, consistent with the prospect theory (Kahneman & Tversky, 1979).

Furthermore, the analysis proceeds with analyzing control variables as, in contrast to earlier models, Models 8 and 9 exhibit significant deviations in the coefficients of the control variables. Model 8 shows a statistically significant negative coefficient (-0.004) for the debt ratio at the 1% level, hence indicating that a one unit increase in a firm's debt ratio is associated with a 0.004-unit decrease in their investment activities. However, Model 9 displays a lower coefficient (-0.000), indistinguishable from zero, which loses its statistical significance. Moreover, market-to-book is statistically significant at the 1% level with a positive coefficient (0.006) pre-Covid. The magnitude of the coefficient (0.002) decreases, and the variable loses its statistical significance after the pandemic onset, possibly due to lower market valuations.

The unprecedented uncertainties of Covid, amplified by extensive media exposure, could have skewed managers' perceptions of risk. During the substantially increased risks of potential revenue and market share losses, managers might have emphasized upholding operations instead of conventional caution regarding debt accrual. Near-zero interest rates presumably facilitated this shift as they effectively established debt as an essential and financially viable instrument for investments to mitigate these adverse effects. Consistent with the prospect theory, this shift in managers' perspective possibly delegated the risks associated with increased debt to a secondary role, which may explain the insignificant debt ratio coefficient in Model 9.

[Table 10](#) further reveals that size is statistically significant at the 1% level through both periods, exhibiting a negative correlation with investment activities. An inverse trend for firm age is also observed. Before the pandemic, the statistically insignificant negative coefficient (-0.007) indicates that firm age was negligible in firms' investment decisions. However, in the Covid era, the firm age coefficient changes direction and substantially increases to 0.158, statistically significant at the 1% level. This implies that older firms, on average, invested more than younger

firms during the pandemic, thus magnifying the relevance of age as an important determinant of firms' investment activities when facing great uncertainty. Older and more well-established firms, leveraging their accumulated experience, may have effectively coped with these uncertainties, potentially explaining the observation of a positive and statistically significant coefficient post-Covid. Consequently, this might further indicate a shift in investors' preferences, demonstrating a bias towards these firms, as these may be perceived as lower risk.

7.2.3 Hypothesis 3

Our third hypothesis, aimed to determine whether there is a divergence in the strength of the predictive efficacy between the FEARS index and market-based measures of investor sentiment, utilized in prior research. For this purpose, we replace the FEARS index with VIX (see also time series in [Figure 6](#)). The results can be found in Model 10 of [Table 11](#). Model 10 reveals a negative coefficient (-0.0575) statistically significant at the 1% level, highlighting a robust inverse correlation between VIX and corporate investments. This implies that a one unit increase in market volatility is associated with, on average, a 0.0575-unit decrease in firms' investment activities. Thus, given that market-based sentiment influences firms' investment behaviors, we find support for our third hypothesis.

Although our results signify a clear difference in strength between the two models, the magnitude of the difference was unexpected. However, it is crucial to consider the fundamentally different approaches of these indices in measuring investor sentiment. Firstly, the VIX index directly captures market signals, specifically those from the volatility of options on the S&P 500 index, which is mainly a reflection of trades by sophisticated market participants (John & Li, 2021). Thus, it mainly represents the sentiment of institutional investors, who, arguably the most informed market participants, have access to highly advanced tools and strategies for effective investment management. Given their resources and capacity, this essentially means that fluctuations in VIX promptly capture shifts in firms' investment behavior. Thus, an increase in the VIX may be interpreted as increased market uncertainty, which may cause firms to delay investment decisions.

Conversely, the FEARS index is more comprehensive as it encompasses the sentiment of a larger demographic, primarily retail investors, who utilize the internet to make investment decisions. However, as retail investors generally have limited resources or lower risk tolerance than sophisticated investors to effectively react to such fluctuations, their sentiment may not directly correlate with market actions.

Consequently, in the context of measuring investor sentiment, the different approaches by these indices result in variances in their relationship with corporate investments. Firms' responses to variations in these indices may vary depending on how they interpret the indications they convey regarding future economic outlooks. Consequently, given that the VIX index may have a more immediate effect. In contrast, the FEARS index impact may be delayed, and the divergence observed in our results to hypothesis 2 may be viewed as an expected outcome, revealing the influence of diverse investor demographics on firms' investment activities.

7.3 Robustness Tests

Our strategy for robustness testing aims to assess the inherent uncertainty in the ideal number of search terms for the FEARS index. Rather than exclusively using thirty search terms, we assess the model's sensitivity against minor alterations in the number of words in the FEARS index. To capture a broad spectrum of configurations, we examine a variety of arbitrarily search term counts, including 25, 20, 15, 10, and 5 words.

[Table 12](#) reveals its sensitivity to the number of words in the FEARS index as Models 4.A to 4.E fail to establish a statistically significant relationship over the sample period, whereas Model 4.F is statistically significant at the 5% level, exhibiting a positive coefficient of 0.018. However, the coefficients in Models 4.A to 4.E are still essentially indistinguishable from zero. Thus, there may be noise in the variable, more specifically, the search terms included in FEARS5 could be more relevant for influencing investment activities. Debt ratio and size remain statistically significant at the 1% level through all configurations of the FEARS index, exhibiting a negative effect on investment activities. Furthermore, market-to-book and board size are statistically

significant at the 1% level throughout the FEARS configurations, exhibiting positive correlations with investment activities, which aligns with the predicted relationships.

[Table 13](#) illustrates how different configurations of the FEARS index influence Model 8, encompassing the pre-Covid period. FEARS is statistically significant at the 1% level for all configurations but FEARS25, with coefficients ranging from 0.023 to 0.0391. Thus, the model does not display any significant sensitivity to different configurations of the FEARS index. As in Model 4, debt ratio and size remain statistically significant at the 1% level through all alterations of the FEARS variable, showing a negative correlation with investment activities. Market-to-book and board size are statistically significant at the 1% level through all configurations, displaying a positive correlation with investment activities. Thus, the relationship between the FEARS index and investment activities is robust regardless of how many search terms are included in the index.

Lastly, [Table 14](#) illustrates how different configurations of the FEARS index influence Model 9, encompassing the Covid period. In contrast to Model 8, FEARS is statistically significant at the 1% level for all configurations, with coefficients ranging from 0.0367 to 0.0604. These results indicate both a more robust and amplified relationship after the onset of the pandemic, which aligns with the notion of the growing influence and power of retail investors. Size and age remain statistically significant at the 1% level through all configurations.

8. Reflection & Discussion

8.1. Model limitations

Despite our attempts to account for numerous factors that may influence corporate investment activities, our model may still inadvertently omit some variables due to their unobservability or measurement difficulties. For example, some firms may have time-varying industry-specific investment opportunities that our model does not fully capture. Thus, this bias results in the model attributing the effect of the missing variables to the included variables.

8.1.1 Discussion of the FEARS-index as a proxy for retail investor sentiment

Throughout this study, we have identified methodological constraints that underscore the complexity of measuring retail investor sentiment. The FEARS index reflects the sentiment of all Google users, not exclusively that of retail investors. Thus, non-investment-related searches could impair the precision of the index. Moreover, the search volume data represents a subset of the population, namely internet users who use Google as their search engine. Thus, this introduces potential selection bias as it excludes the sentiment of investors who gather their financial information elsewhere. Despite these considerations, we argue that the FEARS index remains a meaningful proxy for retail investor sentiment in the U.S., given the high prevalence of American households with stock market investments (Forbes, 2022).

8.2 Discussion in relation to H1

The FEARS index is constructed using search query data over eight years, encompassing two fundamentally different periods in terms of investor sentiment, and as search behaviors evolve, this could affect the validity of the index as a proxy over the full sample period. We hypothesize that 2014 through 2019 was a relatively homogenous period, whereas 2020 through 2021 was a relatively heterogeneous period, during which the onset of the pandemic fundamentally altered the financial landscape, as reflected by the significantly worsened retail investor sentiment ([Table 8](#)).

Upon closer examination of the FEARS index across our sample period ([Figure 5](#)), one peculiar observation is that the index value in the first quarter of 2020 is unaffected by the onset of the pandemic. This may seem counterintuitive initially, given that it was a period of immense global upheaval, and one would expect significant volatility in sentiment. However, we argue that it is reasonable to assume that the search behavior among households changed significantly in this period, where words such as ‘covid,’ ‘pandemic,’ ‘lockdowns,’ and ‘masks’ were among the most entered terms in Google Search. These search terms are not included in the FEARS index, so they do not contribute to its measurement during this crucial period. Nevertheless, including terms such as ‘covid’ or ‘pandemic’ in an updated wordlist would not have changed the index, as these words, despite their relevance to the timeframe, would still have been omitted during the

index construction due to their lack of inherent economic meaning. While these terms lack economic relevance, they have undeniably affected retail investor sentiment and capital market dynamics. If sentiment shifts are triggered by non-economic terms that nonetheless heavily impact capital markets, the FEARS index fails to portray the change in sentiment accurately. This consideration raises questions about the efficacy of the FEARS index in the face of unprecedented events such as the pandemic. Thus, we see the potential for an enhanced version of the index. By allowing for dynamic wordlists that adapt over time, the index could become more responsive to current sentiment and shifts in focus, effectively becoming a more sensitive barometer of prevailing attitudes. For instance, in our current financial landscape, terms such as debt ceiling might be relevant to include, reflecting immediate economic concerns.

8.3 Discussion in relation to H2 and endogeneity in the model

Although the study utilizes a fixed effects approach, it does not address endogeneity associated with the correlation between FEARS and the error term (Robert & Whited, 2013). Thus, we acknowledge a potential endogeneity issue with the FEARS index that merits close attention as it may introduce biases into our findings. The index may concurrently be influenced by broader economic conditions, which complicates our ability to isolate the effects of retail investor sentiment on corporate investment policies, rendering the interpretation of our findings more challenging. However, Da et al. (2015) argue that even if the FEARS index is endogenous to macroeconomic events, it does not invalidate the use of search as a measure of sentiment, suggesting that the FEARS index is anticipated to be endogenous to macroeconomic events.

While our study identifies a correlation between the FEARS index and corporate investment activities in our second hypothesis, it is imperative to underscore that this does not imply causation. Thus, our findings must be interpreted cautiously, acknowledging that our model may capture correlative rather than causative relationships. This identification problem makes reliable inference practically impossible. Our general hypothesis posits that retail investor sentiment can influence investment policies. However, it is entirely plausible that changes in investment activities could, in turn, affect retail investor sentiment as we cannot infer the nature of the relationship and thereby not conclude whether firms cater to retail investors' "animal spirits".

Consequently, should a causative relationship exist, changes in investor sentiment could be both a precursor and an outcome of firms' investment policy decisions, thereby introducing a potential reverse causality issue.

9. Conclusion

Prior research has established a positive relationship between investor sentiment and corporate investments (Baker & Wurgler, 2007; Polk & Sapienza, 2009; Zhu, Song & Huang, 2018; Du & Hu, 2019; Li & Grundy, 2023). However, traditional measures of investor sentiment often fail to differentiate between retail and institutional investors. Given the rapidly increasing retail investor participation in the stock market, this study aims to investigate retail investors' growing influence in shaping corporate investment decisions by utilizing an innovative measure that allows us to isolate retail investor sentiment. It further seeks to understand whether firms engage in a balancing act between adhering to traditional financial principles aimed at maximizing shareholder value and catering to the market sentiment, often swayed by retail investors' collective "animal spirits".

Our empirical analysis draws upon a strongly-balanced panel of 741 publicly listed firms on the Nasdaq Composite from 2014 to 2021. Utilizing multiple linear regression models, including fixed effects and clustered standard errors, our initial analysis did not reveal a statistically significant association between retail investor sentiment and firms' investment activities across the full sample period. We hypothesized that 2014 through 2019 was a relatively homogenous period, whereas investor sentiment negatively worsened during the pandemic. In the pre-Covid and Covid subsample analysis, we discovered a robust correlation between retail investor sentiment and investment activities. Further, this study finds an inverse relationship between the FEARS index and VIX, suggesting that retail and institutional investors may influence corporate investment decisions differently.

This study utilizes the FEARS index to effectively gauge investor sentiment and its effects on firms' investment activities, offering a more granular representation of investor sentiment that conventional market-based measures may not capture. Thus, this study extends the

understanding of sentiment dynamics by focusing on a broad spectrum of investors, primarily retail investors, an aspect that prior research generally neglects. We further analyze positive and negative sentiment, contributing greater nuance to prior research that primarily considers negative sentiment. However, we discovered varying model sensitivity to the number of search terms and an unexpected stability of the index during the pandemic's onset, underlining the potential need for dynamic wordlists. A limitation of our study is the difficulty of inferring our results due to the potential endogeneity issue with the FEARS index. Thus, we cannot conclude whether firms cater to retail investors' "animal spirits". We further acknowledge the possibility that the model may inadvertently omit variables.

This study may be further extended for future research by utilizing a dynamic wordlist to construct the FEARS index. Given that the relevance and usage of search terms can evolve over time, a dynamic wordlist may improve the accuracy and relevance of the sentiment measure. Likewise, an interesting avenue could be exploring the influence of only positive terms, discerning whether investor optimism has a different influence on corporate investment policies than overall sentiment. Further, it would be interesting to include additional controls, such as GDP, as investment opportunities in many industries are a function of the business cycle.

Reference list

- Agnes Cheng, C. S., Kite, D., & Radtke, R. (1994). The applicability and usage of NPV and IRR capital budgeting techniques. *Managerial Finance*, 20(7), 10-36.
- Aivazian, V. A., Ge, Y., & Qiu, J. (2005). The impact of leverage on firm investment: Canadian evidence. *Journal of corporate finance*, 11(1-2), 277-291.
- Al-Thaqeb, S. A., Algharabali, B. G., & Alabdulghafour, K. T. (2022). The pandemic and economic policy uncertainty. *International Journal of Finance & Economics*, 27(3), 2784-2794.
- Arif, S., & Lee, C. M. (2014). Aggregate investment and investor sentiment. *The Review of Financial Studies*, 27(11), 3241-3279.
- Arora, K. (2022). *Council Post: The Rise Of The Retail Investor*. Forbes. Available at: <https://www.forbes.com/sites/forbesagencycouncil/2022/11/04/the-rise-of-the-retail-investor/?sh=7db898bf1755> [Accessed: 2023-04-23]
- Audrino, F., Sigrist, F., & Ballinari, D. (2020). The impact of sentiment and attention measures on stock market volatility. *International Journal of Forecasting*, 36(2), 334-357.
- Baker, M., & Wurgler, J. (2002). Market timing and capital structure. *The journal of finance*, 57(1), 1-32.
- Baker, M., & Wurgler, J. (2004). A catering theory of dividends. *The Journal of finance*, 59(3), 1125-1165.
- Baker, M., & Wurgler, J. (2006). Investor sentiment and the cross-section of stock returns. *The Journal of Finance*, 61(4), 1645-1680.

Baker, M., & Wurgler, J. (2007). Investor sentiment in the stock market. *Journal of economic perspectives*, 21(2), 129-151.

Baker, M., & Wurgler, J. (2013). Behavioral corporate finance: An updated survey. In *Handbook of the Economics of Finance*(Vol. 2, pp. 357-424). Elsevier.

Baker, M., Stein, J. C., & Wurgler, J. (2003). When does the market matter? Stock prices and the investment of equity-dependent firms. *Quarterly Journal of Economics*, 118(3), 969-1005.

Barberis, N., Shleifer, A., & Vishny, R. (1998). A model of investor sentiment. *Journal of financial economics*, 49(3), 307-343.

Baysinger, B. D., & Butler, H. N. (2019). Corporate governance and the board of directors: Performance effects of changes in board composition. In *Corporate Governance* (pp. 215-238). Gower.

Bertrand, M., & Schoar, A. (2003). Managing with style: The effect of managers on firm policies. *The Quarterly journal of economics*, 118(4), 1169-1208.

Bikhchandani, S., Hirshleifer, D., & Welch, I. (1992). A theory of fads, fashion, custom, and cultural change as informational cascades. *Journal of political Economy*, 100(5), 992-1026.

Bilel, H., & Mondher, K. (2021). What Can explain catering of dividend? Environment information and investor sentiment. *Journal of Economics and Finance*, 45, 428-450.

Boone, A. L., Field, L. C., Karpoff, J. M., & Raheja, C. G. (2007). The determinants of corporate board size and composition: An empirical analysis. *Journal of financial Economics*, 85(1), 66-101.

Bouri, E., Gupta, R., & Roubaud, D. (2019). Herding behaviour in cryptocurrencies. *Finance Research Letters*, 29, 216-221.

Brown, G. W., & Cliff, M. T. (2005). Investor sentiment and asset valuation. *The Journal of Business*, 78(2), 405-440.

Baxamusa, M. (2011). How well do market timing, catering, and classical theories explain corporate decisions?. *Journal of Financial Research*, 34(2), 217-239.

Chao, C. H., Huang, C. J., Ho, R. J., & Huang, H. Y. (2022). Catering to investors through capital expenditures: Testing assets substitution problem around financing. *The North American Journal of Economics and Finance*, 59, 101561.

Chazi, A., Theodossiou, A., & Zantout, Z. (2018). Corporate payout-form: investors' preference and catering theory. *Managerial Finance*.

Chen, C. J., & Jaggi, B. (2000). Association between independent non-executive directors, family control and financial disclosures in Hong Kong. *Journal of Accounting and Public policy*, 19(4-5), 285-310.

Cleary, S. (1999). The relationship between firm investment and financial status. *The journal of finance*, 54(2), 673-692.

Cronqvist, H., Makhija, A. K., & Yonker, S. E. (2012). Behavioral consistency in corporate finance: CEO personal and corporate leverage. *Journal of financial economics*, 103(1), 20-40.

Da, Z., Engelberg, J., & Gao, P. (2015). The sum of all FEARS investor sentiment and asset prices. *The Review of Financial Studies*, 28(1), 1-32.

Deloitte. The Future of Retail Brokerage. (2021). Deloitte United States. Available at: <https://www2.deloitte.com/us/en/pages/financial-services/articles/the-future-of-retail-brokerage.html> [Accessed: 2023-05-16]

Denis, D. J. (2015). Doing empirical corporate finance research in an endogenous world: Tutorial session from the 2014 FMA annual meeting. *Journal of Applied Finance (Formerly Financial Practice and Education)*, 25(1).

Du, D., & Hu, O. (2019). Why does stock-market investor sentiment influence corporate investment?. *Review of Quantitative Finance and Accounting*, 54(4), 1221-1246.

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

Fazzari, S., Hubbard, R. G., & Petersen, B. C. (1987). Financing constraints and corporate investment.

FAQ about Google Trends data - Trends Help. (2023). Support.google.com. Available at: https://support.google.com/trends/answer/4365533?hl=en&ref_topic=6248052&sjid=3530489346884068025-EU

Graham, J. R., & Harvey, C. R. (2001). The theory and practice of corporate finance: Evidence from the field. *Journal of financial economics*, 60(2-3), 187-243.

Griffin, D., & Tversky, A. (1992). The weighing of evidence and the determinants of confidence. *Cognitive psychology*, 24(3), 411-435.

Gu, L., & Zhang, H. (2022). Stakeholder-oriented corporate investment: A catering perspective. *Finance Research Letters*, 48, 102863.

Guiso, L., Sapienza, P., & Zingales, L. (2008). Trusting the stock market. *The Journal of Finance*, 63(6), 2557-2600.

Hirshleifer, D., & Hong Teoh, S. (2003). Herd behaviour and cascading in capital markets: A review and synthesis. *European Financial Management*, 9(1), 25-66.

IOSCO. (2022). *Report on Retail Distribution and Digitalisation Consultation Report The Board of the International Organization of Securities Commissions*. Available at: <https://www.iosco.org/library/pubdocs/pdf/IOSCOPD695.pdf> [Accessed: 2023-04-20]

Jensen, M. C. (1986). Agency costs of free cash flow, corporate finance, and takeovers. *The American economic review*, 76(2), 323-329.

Jensen, M. C. (1993). The modern industrial revolution, exit, and the failure of internal control systems. *The Journal of Finance*, 48(3), 831-880.

Jiang, F., Lee, J., Martin, X., & Zhou, G. (2019). Manager sentiment and stock returns. *Journal of Financial Economics*, 132(1), 126-149.

John, K., & Li, J. (2021). COVID-19, volatility dynamics, and sentiment trading. *Journal of Banking & Finance*, 133, 106162.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2), 363-391.

Kahneman, D. (2003). Maps of bounded rationality: Psychology for behavioral economics. *American economic review*, 93(5), 1449-1475.

Kayhan, A., & Titman, S. (2007). Firms' histories and their capital structures. *Journal of financial Economics*, 83(1), 1-32.

Keynes, J. M. (1937). The general theory of employment. *The quarterly journal of economics*, 51(2), 209-223.

Kusnadi, Y., & Wei, K. J. (2017). The equity-financing channel, the catering channel, and corporate investment: International evidence. *Journal of Corporate Finance*, 47, 236-252.

Lang, L., Ofek, E., & Stulz, R. (1996). Leverage, investment, and firm growth. *Journal of financial Economics*, 40(1), 3-29.

Lee, I., & Loughran, T. (1998). Performance following convertible bond issuance. *Journal of Corporate Finance*, 4(2), 185-207.

Liao, S., & Errico, M. (2022). Corporate investment and stock market valuation. *Journal of Business Finance & Accounting*.

Li, H., & Grundy, B. (2010). Investor sentiment, executive compensation, and corporate investment. *Journal of Banking & Finance*, 34(10), 2439-2449.

Li, H., & Grundy, B. (2023). The effect of investor sentiment and the structure of shareholder ownership on corporate investment. *International Journal of Managerial Finance*, 19(1), 155-172.

Lobo, G. J. (2017). Accounting research in banking: A review. *China Journal of Accounting Research*, 10(1), 1-7.

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

Loughran, T., & McDonald, B. (2020). Textual analysis in finance. *Annual Review of Financial Economics*, 12, 357-375.

Modigliani & Miller (1958). *The Cost of Capital, Corporation Finance and the Theory of Investment*.

Myers, S. C. (1977). Determinants of corporate borrowing. *Journal of financial economics*, 5(2), 147-175.

Myers, S. C. (1984). Capital structure puzzle.

Myers, S. C., & Majluf, N. S. (1984). Corporate financing and investment decisions when firms have information that investors do not have. *Journal of financial economics*, 13(2), 187-221.

Neves, M. E. D. (2017). Payout and Firm's Catering. *International Journal of Economics & Business Administration (IJEBA)*, (4), 104-132.

Osipovich, A. (2020). Individual-Investor Boom Reshapes U.S. Stock Market. *The Wall Street Journal (WSJ)*. Available at:
<https://www.wsj.com/articles/individual-investor-boom-reshapes-u-s-stock-market-11598866200>
[Accessed: 2023-05-16]

Penman, S. H. (1996). The articulation of price-earnings ratios and market-to-book ratios and the evaluation of growth. *Journal of accounting research*, 34(2), 235-259.

Polk, C., & Sapienza, P. (2002). The real effects of investor sentiment. *Available at SSRN* 354386.

Polk, C., & Sapienza, P. (2004). The real effects of investor sentiment.

Polk, C., & Sapienza, P. (2009). The stock market and corporate investment: A test of catering theory. *The Review of Financial Studies*, 22(1), 187-217.

Rajan, R. G., & Zingales, L. (1995). What do we know about capital structure? Some evidence from international data. *The Journal of Finance*, 50(5), 1421-1460.

Rajgopal, S., Shivakumar, L., & Simpson, A. V. (2007). A catering theory of earnings management. *Available at SSRN 991138*.

Roberts, M. R., & Whited, T. M. (2013). Endogeneity in empirical corporate finance¹. In *Handbook of the Economics of Finance* (Vol. 2, pp. 493-572). Elsevier.

Scharfstein, D. S., & Stein, J. C. (1990). Herd behavior and investment. *The American economic review*, 465-479.

Simon, H. A. (1955). A behavioral model of rational choice. *The quarterly journal of economics*, 99-118.

Singer, E. (2002). The use of incentives to reduce nonresponse in household surveys. *Survey nonresponse*, 51(1), 163-177.

Tetlock, P. C. (2007). Giving content to investor sentiment: The role of media in the stock market. *The Journal of finance*, 62(3), 1139-1168.

Tetlock, P. C., Saar-Tsechansky, M., & Macskassy, S. (2008). More than words: Quantifying language to measure firms' fundamentals. *The journal of finance*, 63(3), 1437-1467.

The General Inquirer. (n.d.). *Spreadsheet Guide*. Available at:

http://www.wjh.harvard.edu/~inquirer/spreadsheet_guide.htm [Accessed: 2023-04-05]

U.S. Census Bureau. (n.d.). North American Industry Classification System (NAICS). Available at: <https://www.census.gov/naics/?58967?yearbck=2022> [Accessed: 2023-04-25]

U.S. Securities and Exchange Commission. (2023). Exchange Act reporting requirements. Available at: <https://www.sec.gov/education/smallbusiness/goingpublic/exchangeactreporting> [Accessed: 2023-04-21]

Zhu, Z., Song, X., & Huang, W. (2018). Top management team demography, investor sentiment, and the investment levels of listed companies. *Asia-Pacific Journal of Accounting & Economics*, 25(3-4), 481-495.

Figures

Figure 1 - Search volume index (SVI) change for the search term 'Silver'

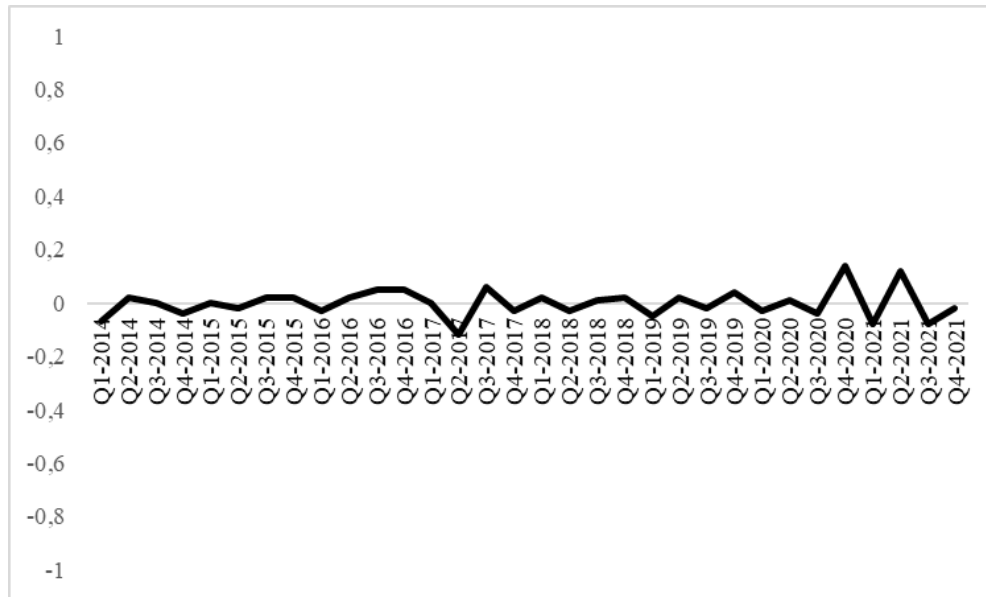


Figure 2 - Search volume index (SVI) change for the search term 'Charitable Donation'

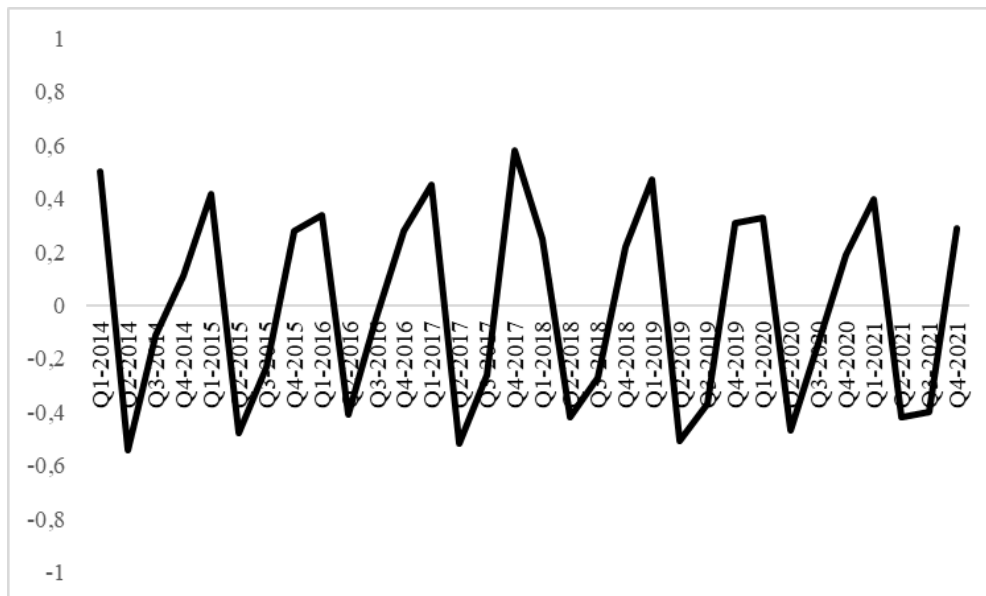


Figure 3 - Histogram for t-statistics on adjusted quarterly change in search volume index (ASVI) for all search terms, excluding tails

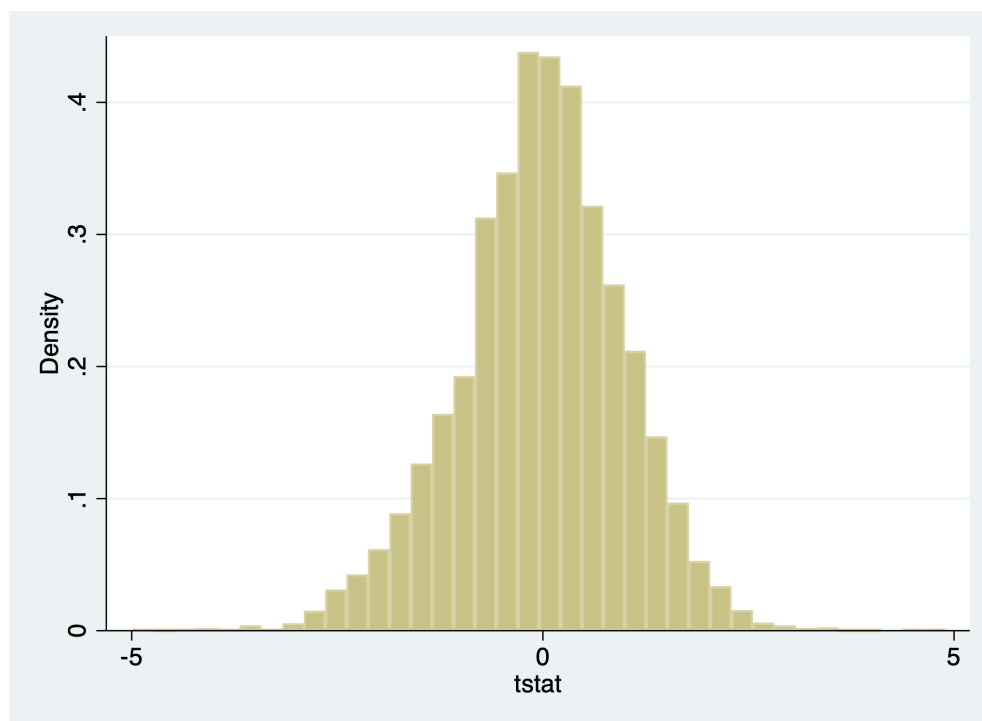


Figure 4 - Histogram for t-statistics on adjusted quarterly change in search volume index (ASVI) for all search terms, including tails

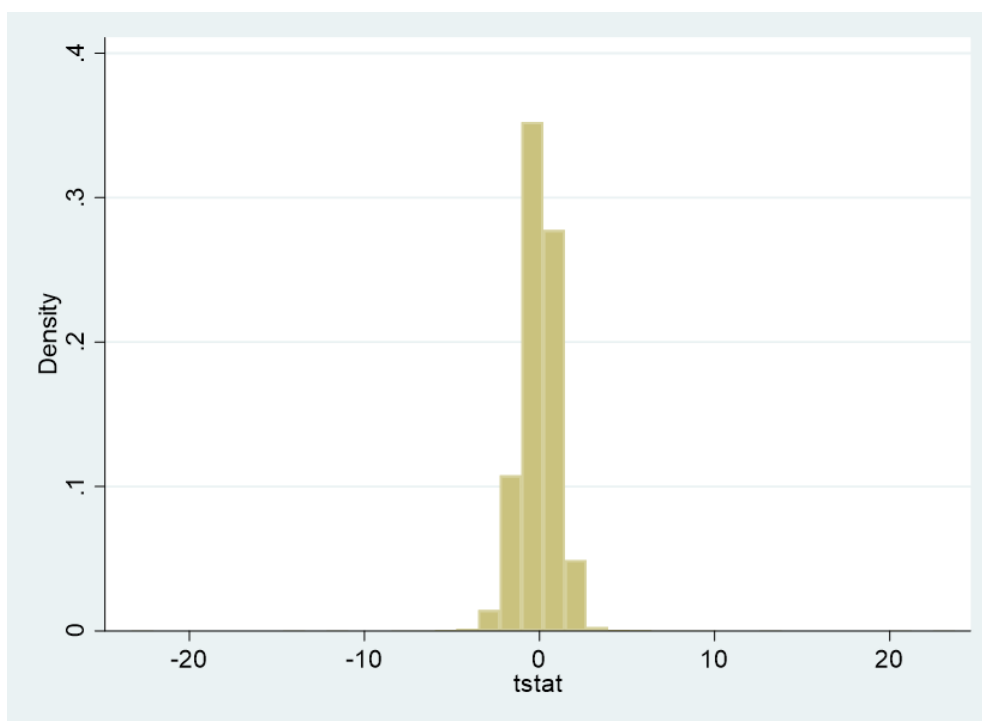


Figure 5 - The FEARS30 Index Variable

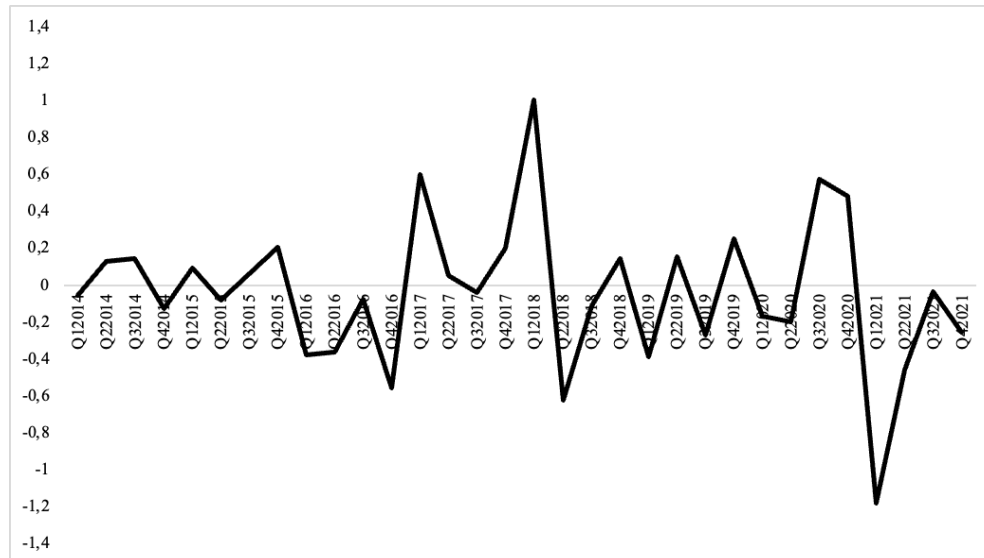
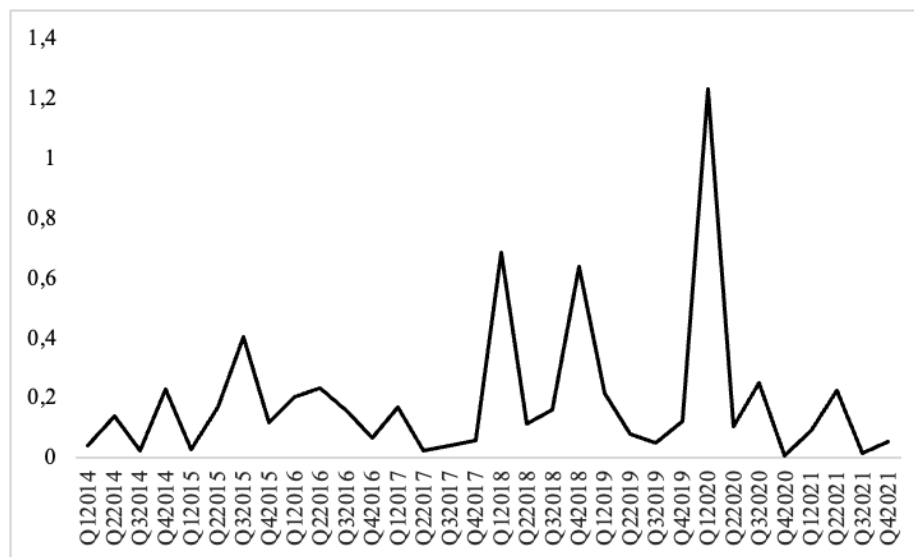


Figure 6 - The VIX Index Variable



Tables

Table 1 - Summary statistics

Table 1 reports summary statistics on 23712 firm-quarter observations (IA) representing 741 individual firms on the Nasdaq Composite (CCMP) over the sample period 2014-2021. All accounting-related variables (FCF, DR, TA, ROA, MB) have been winsorized at the top and bottom 1%. Investment activities is defined as the average capital expenditures and R&D spending over the next three quarters divided by total assets. FEARS is the proxy for retail investor sentiment, and it represents the thirty search terms with the strongest relationship with CCMP market returns for quarter t. Free cash flow is defined as the twelve trailing months' cash flow from operating activities divided by total assets in the quarter prior to the twelve trailing month period. The debt ratio is measured as total debt relative to total assets. Return on assets is measured as the net income divided by total assets. Market-to-book is calculated as the equity market value divided by the equity book value. Board size measures the number of directors on the board, whereas board independence measures the proportion of independent directors on the board. Chair duality is a dummy variable assigned a value of 1 if the CEO is also the chair of the board and zero otherwise.

Variable	Symbol	Mean	Median	SD	Min	Max	N
Dependent variable							
Investment Activities	IA	.01	.01	0.01	0	.06	23712
Main Explanatory Variable							
Financial and Economic Attitudes Revealed by Search	FEARS30	-.04	-.05	0.40	-1.18	1	23744
Control Variables							
Free Cash Flow	FCF	.01	.08	0.29	-1.4	.43	23739
Debt Ratio	DR	.22	.16	0.24	0	1.25	23743
Total Assets	TA	3563.69	486.57	10846.89	7.23	77093	23743
Return on Assets	ROA	-.06	.03	0.30	-1.49	.37	23712
Market-to-Book	MB	4.09	2.51	7.60	-22.65	46.96	23743
Firm Age	Age	39.78	32	28.13	0	171	23744
Board Size	BoardSize	7.95	8	2.04	2	18	23744
Board Independence	Ind	.82	.86	0.09	0	1	23744
Chair Duality	Dual	.34	0	0.47	0	1	23744

Table 2 - Descriptive statistics of the t-statistics from the regression of $\Delta ASVI$ on market returns for all the search terms

Table 2 reports the descriptive statistics of the t-statistics obtained from regressing the adjusted search volume index (ASVI) on CCMP market returns for all our search terms. The resulting distribution exhibits relative symmetry with a slight skew towards the right. Moreover, significant outliers can be observed in both tails of the distribution. A larger absolute value of the t-statistics signifies a stronger relationship between the corresponding search term and market returns.

Variables	Obs	Mean	Std. Dev.	Min	Max	p1	p99	Skew.	Kurt.
T-statistic	7683	-.023	1.183	-23.3	23.724	-2.641	2.325	.56	62.791

Table 3 - Pairwise correlation matrix

Table 3 presents a pairwise correlation matrix, illustrating the interrelationships between all independent variables used in the multivariate regression model. Each cell in the table shows the correlation coefficient between two variables, ranging from -1 to 1. A correlation of 1 indicates a perfect positive correlation, where the two variables move in the same direction proportionally. A correlation of -1 indicates a perfect negative correlation, where the two variables move in opposite directions. A correlation coefficient of 0 suggests no linear relationship between the two variables. The diagonal from the top left to bottom right contains the number 1, because each variable has a perfect positive correlation with itself.

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
(1) FEARS30	1.000									
(2) FCF	-0.007	1.000								
(3) DR	-0.021	-0.011	1.000							
(4) Size	-0.018	0.477	0.221	1.000						
(5) ROA	-0.012	0.851	-0.061	0.488	1.000					
(6) MB	-0.012	-0.011	-0.073	0.085	-0.023	1.000				
(7) Age	-0.013	0.203	0.065	0.200	0.235	-0.092	1.000			
(8) BoardSize	-0.013	0.206	0.186	0.633	0.214	0.101	0.196	1.000		
(9) Ind	-0.014	0.072	0.096	0.273	0.075	0.052	0.046	0.331	1.000	
(10) Dual	0.002	0.045	-0.002	-0.017	0.046	0.003	-0.002	-0.099	-0.189	1.000

Table 4 - Variance inflation factor

Table 4 displays the variance inflation factor (VIF) for each variable included in our multivariate regression model. The VIF is a measure that quantifies how much the variance of the estimated regression coefficients is increased due to multicollinearity. A VIF of 1 suggests that there is no correlation between the given independent variable and the others. Higher VIF values indicate a higher degree of correlation and potential multicollinearity. A common rule of thumb is that a VIF greater than 5 might indicate problematic multicollinearity.

Variable	VIF	1/VIF
ROA	3.895	.257
FCF	3.706	.270
Size	2.298	.435
BoardSize	1.823	.548
Ind	1.169	.855
DR	1.115	.897
Age	1.099	.910
Dual	1.047	.955
MB	1.041	.961
FEARS30	1.001	.999
Mean VIF	1.819	-

Table 5 - Test of Heteroscedasticity

Table 5 presents the results of the White's test utilized to detect heteroskedasticity in the error terms of our multivariate regression model. The null hypothesis of the White test is homoscedasticity, meaning that the variance of the error terms is constant. Based on the p-value of 0.000, the null hypothesis is rejected and we infer that heteroskedasticity is present in the data.

White test of heteroscedasticity	Chi2 statistic	df	p-value
H_0 : Homoscedasticity	1236.63	64	0.000

Table 6 - The Hausman (1978) specification test

Table 6 presents the results of the Hausman specification test, used to determine whether a random effects model or a fixed effects model is more appropriate. The null hypothesis posits that the random effects estimator is preferable. The alternative hypothesis posits that the fixed effects estimator should be used due to potential correlation between the unobserved effects and the explanatory variables. Based on the p-value of 0.000, we reject the null hypothesis and conclude that the fixed effects model is more appropriate.

χ^2	407.259
p-value	0.000

Table 7 - Regression results from different model specifications

Table 7 displays the outputs from four distinct model specifications: Model 1 with default standard errors and no fixed effects, Model 2 with robust standard errors and no fixed effects, Model 3 with standard errors clustered by firm and no fixed effects, and Model 4 with standard errors clustered by firm and fixed effects. The table provides a side-by-side comparison of these models based on the specifications discussed in the pre-regression diagnostics. It reveals that the coefficients remain consistent across the first three model specifications (1-3), as expected. Model 4, which includes both fixed effects and clustered standard errors, is identified as the preferred specification for subsequent hypothesis testing. This table, therefore, serves as a comprehensive guide to the regression results under different model specifications, offering insights into the stability and robustness of the model outcomes.

	Model 1	Model 2	Model 3	Model 4
Dep. Variable	Investment Activities	Investment Activities	Investment Activities	Investment Activities
FEARS30	0.000143 (0.000119)	0.000143* (8.26e-05)	0.000143* (8.26e-05)	-1.58e-06 (8.24e-05)
FCF	0.00153*** (0.000380)	0.00153* (0.000822)	0.00153* (0.000822)	0.000721 (0.000856)
DR	-5.13e-05*** (3.69e-06)	-5.13e-05*** (1.07e-05)	-5.13e-05*** (1.07e-05)	-5.51e-05*** (1.19e-05)
Size	-0.00150*** (0.000104)	-0.00150*** (0.000249)	-0.00150*** (0.000249)	-0.00254*** (0.000376)
ROA	1.14e-05*** (3.74e-06)	1.14e-05 (8.47e-06)	1.14e-05 (8.47e-06)	1.45e-05* (8.72e-06)
MB	6.92e-05*** (7.67e-06)	6.92e-05*** (1.54e-05)	6.92e-05*** (1.54e-05)	6.80e-05*** (1.53e-05)
Age	-7.68e-06 (9.43e-06)	-7.68e-06 (1.27e-05)	-7.68e-06 (1.27e-05)	-6.08e-05 (6.71e-05)
Board Size	0.000407*** (5.62e-05)	0.000407*** (0.000131)	0.000407*** (0.000131)	0.000392*** (0.000138)
Independence	-4.82e-05*** (9.39e-06)	-4.82e-05** (2.13e-05)	-4.82e-05** (2.13e-05)	-3.72e-05* (2.21e-05)
Dual	0.000735*** (0.000189)	0.000735 (0.000562)	0.000735 (0.000562)	0.000684 (0.000610)
Constant	0.0203*** (0.000991)	0.0203*** (0.00225)	0.0203*** (0.00225)	0.0283*** (0.00321)
Observations	23,712	23,712	23,712	23,712
R-squared				0.043
Standard Errors	Default	Robust	Clustered (firm)	Clustered (firm)
Fixed Effects	No	No	No	Yes
Number of id	741	741	741	741

Table 8 - Tests of differences in means

Table 8 comprehensively analyzes the differences in means for variables, including Investment Activities, FEARS30, and VIX. The table displays sample size (N) and the means for both the pre-Covid and post-Covid periods. It also provides the population size (N), revealing the dataset scale from which the sample is drawn. The difference between the pre-Covid and post-Covid means (0-1) is calculated and shown alongside the associated p-value, indicating the observed difference's statistical significance.

Variables	Pre-Covid		Post-Covid		Difference (0-1)	p-value
	N	Mean	N	Mean		
Investment Activities	17784	0.009	5928	0.008	-0,001	0.0000
FEARS30	17808	-0.002	5936	-0.15	0.154	0.0000

Table 9 - Regression results including quarter dummy variables

Table 9 presents the estimation results of the primary regression Model 4. This model incorporates firm fixed effects and uses clustered standard errors, providing a more nuanced perspective on the dataset. The table showcases how these elements interact within the model, offering insights into the implications of these specific methodological choices on the results.

	Model 4	Model 5	Model 6	Model 7
Period	2014Q1 - 2021Q4	2014Q1 - 2021Q4 excl Q12018	2014Q1 - 2021Q4 excl Q12021	2014Q1 - 2021Q4 excl Q12018 & Q12021
Dep. Variable	Investment Activities	Investment Activities	Investment Activities	Investment Activities
FEARS30	-1.58e-06 (8.24e-05)	-0.000282*** (8.71e-05)	0.000159* (9.47e-05)	-0.000168 (0.000124)
FCF	0.000721 (0.000856)	0.000743 (0.000857)	0.000717 (0.000856)	0.000739 (0.000858)
DR	-0.00551*** (0.00119)	-0.00546*** (0.00119)	-0.00549*** (0.00119)	-0.00545*** (0.00119)
Size	-0.00254*** (0.000376)	-0.00253*** (0.000376)	-0.00254*** (0.000376)	-0.00253*** (0.000376)
ROA	0.00145* (0.000872)	0.00144* (0.000872)	0.00146* (0.000872)	0.00144* (0.000872)
MB	6.80e-05*** (1.53e-05)	6.81e-05*** (1.53e-05)	6.76e-05*** (1.53e-05)	6.79e-05*** (1.53e-05)
Age	-6.08e-05 (6.71e-05)	-7.43e-05 (6.73e-05)	-7.27e-05 (6.72e-05)	-8.00e-05 (6.72e-05)
BoardSize	0.000392*** (0.000138)	0.000392*** (0.000138)	0.000392*** (0.000138)	0.000391*** (0.000138)
Ind	-0.00372* (0.00221)	-0.00373* (0.00221)	-0.00372* (0.00221)	-0.00373* (0.00221)
Dual	0.000684 (0.000610)	0.000690 (0.000609)	0.000683 (0.000610)	0.000689 (0.000610)
dummy_2018q1	- (0.000297)	0.00132*** (0.000297)	- (0.000245)	0.00122*** (0.000332)
dummy_2021q1	- (0.000285)	- (0.000285)	0.000763*** (0.000245)	0.000439 (0.000285)
Constant	0.0283*** (0.00321)	0.0287*** (0.00322)	0.0287*** (0.00321)	0.0289*** (0.00321)
Observations	23,712	23,712	23,712	23,712
R-squared	0.043	0.043	0.043	0.043
Standard Errors	Clustered (firm)	Clustered (firm)	Clustered (firm)	Clustered (firm)
Number of id	741	741	741	741

Table 10 - Regression results for pre-Covid and under Covid

Table 10 displays the estimation results from Models 8 and 9 segmented into pre-Covid and Covid periods. By doing this, the table compares the model outcomes in distinctly different environments. This separation offers insights into the model's performance under varying conditions, facilitating a comprehensive understanding of the model's robustness and adaptability during both periods.

	Model 8	Model 9
Period	2014Q1 - 2019Q4	2020Q1 - 2021Q4
Dep. Variable	Investment Activities	Investment Activities
FEARS30	0.000273*** (0.000102)	0.000520*** (0.000139)
FCF	0.000547 (0.000972)	0.000998 (0.00172)
DR	-3.78e-05*** (1.24e-05)	-4.42e-06 (2.14e-05)
Size	-0.00339*** (0.000432)	-0.00384*** (0.000997)
ROA	1.85e-05* (1.02e-05)	-3.39e-06 (1.33e-05)
MB	5.89e-05*** (1.63e-05)	1.72e-05 (2.08e-05)
Age	-6.63e-05 (8.16e-05)	0.00158*** (0.000302)
Board Size	0.000305** (0.000155)	-8.11e-05 (0.000248)
Independence	-3.94e-05 (2.50e-05)	-8.12e-06 (3.38e-05)
Dual	0.00107 (0.000656)	-0.000211 (0.000561)
Constant	0.0341*** (0.00399)	-0.0330** (0.0129)
Observations	17,784	5928
R-squared	0.041	0.033
Standard Errors	Clustered (firm)	Clustered (firm)
Number of ID	741	741

Table 11 - Regression results including the VIX index

Table 11 presents the impact on regression outcomes when the FEARS index is replaced with the VIX index. This change is visualized in Figure 6's time series, and the table presents the resulting shifts in regression outcomes. This allows for an analysis of the relative influence of the VIX index on the model, providing an alternative perspective to the previous analysis with the FEARS index.

	Model 4	Model 10
Dep. Variable	Investment Activities	Investment Activities
FEARS30	-1.58e-06 (8.24e-05)	- -
VIX	- -	-0.000575*** (0.000103)
FCF	0.000721 (0.000856)	0.000702 (0.000856)
DR	-0.00551*** (0.00119)	-0.00550*** (0.00119)
Size	-0.00254*** (0.000376)	-0.00253*** (0.000376)
ROA	0.00145* (0.000872)	0.00145* (0.000872)
MB	6.80e-05*** (1.53e-05)	6.83e-05*** (1.54e-05)
Age	-6.08e-05 (6.71e-05)	-5.93e-05 (6.66e-05)
Board Size	0.000392*** (0.000138)	0.000391*** (0.000138)
Independence	-0.00372* (0.00221)	-0.00373* (0.00221)
Dual	0.000684 (0.000610)	0.000684 (0.000609)
Constant	0.0283*** (0.00321)	0.0282*** (0.00319)
Observations	23,712	23,712
R-squared	0.043	0.043
Standard Errors	Clustered (firm)	Clustered (firm)
Number of id	741	741

Table 12 - Robustness tests of Model 4

Table 12 demonstrates the impact of various configurations of the FEARS index on Model 4. By testing different search term counts, including 25, 20, 15, 10, and 5 words, this table encapsulates a broad spectrum of potential configurations, offering robustness to the model. The data set encompasses the full sample period and is used specifically to test the first hypothesis of our study. This table allows readers to understand how the model's output changes with different FEARS index configurations.

	Model 4.A	Model 4.B	Model 4.C	Model 4.D	Model 4.E	Model 4.F
Dep. Variable	Investment Activities	Investment Activities	Investment Activities	Investment Activities	Investment Activities	Investment Activities
	FEARS30	FEARS25	FEARS20	FEARS15	FEARS10	FEARS5
FEARS	-1.58e-06 (8.24e-05)	-2.46e-05 (7.96e-05)	8.04e-05 (8.39e-05)	2.97e-05 (6.99e-05)	1.19e-06 (6.57e-05)	0.000182** (8.85e-05)
FCF	0.000721 (0.000856)	0.000721 (0.000856)	0.000721 (0.000857)	0.000721 (0.000857)	0.000721 (0.000857)	0.000728 (0.000857)
DR	-0.00551*** (0.00119)	-0.00551*** (0.00119)	-0.00551*** (0.00119)	-0.00551*** (0.00119)	-0.00551*** (0.00119)	-0.00550*** (0.00119)
Size	-0.00254*** (0.000376)	-0.00254*** (0.000376)	-0.00254*** (0.000376)	-0.00254*** (0.000376)	-0.00254*** (0.000376)	-0.00254*** (0.000376)
ROA	0.00145* (0.000872)	0.00145* (0.000871)	0.00145* (0.000872)	0.00145* (0.000872)	0.00145* (0.000872)	0.00145* (0.000872)
MB	6.80e-05*** (1.53e-05)	6.80e-05*** (1.53e-05)	6.80e-05*** (1.53e-05)	6.80e-05*** (1.53e-05)	6.80e-05*** (1.53e-05)	6.82e-05*** (1.53e-05)
Age	-6.08e-05 (6.71e-05)	-6.14e-05 (6.70e-05)	-5.83e-05 (6.71e-05)	-5.95e-05 (6.72e-05)	-6.07e-05 (6.71e-05)	-4.87e-05 (6.71e-05)
BoardSize	0.000392*** (0.000138)	0.000392*** (0.000138)	0.000393*** (0.000138)	0.000392*** (0.000138)	0.000392*** (0.000138)	0.000393*** (0.000138)
Ind	-0.00372* (0.00221)	-0.00372* (0.00221)	-0.00372* (0.00221)	-0.00372* (0.00221)	-0.00372* (0.00221)	-0.00372* (0.00221)
Dual	0.000684 (0.000610)	0.000684 (0.000610)	0.000686 (0.000610)	0.000685 (0.000610)	0.000685 (0.000610)	0.000691 (0.000610)
Constant	0.0283*** (0.00321)	0.0283*** (0.00321)	0.0282*** (0.00321)	0.0282*** (0.00321)	0.0283*** (0.00321)	0.0278*** (0.00321)
Observations	23,712	23,712	23,712	23,712	23,712	23,712
R-squared	0.043	0.043	0.043	0.043	0.043	0.043
Standard Errors	Clustered (firm)	Clustered (firm)	Clustered (firm)	Clustered (firm)	Clustered (firm)	Clustered (firm)
Number of id	741	741	741	741	741	741

Table 13 - Robustness tests of Model 8

Table 13 presents the effects of varying configurations of the FEARS index on Model 8, focusing specifically on the pre-Covid period (Q1 2014-Q4 2019). It aims to show how changes in the FEARS index parameters can influence the output of Model 8 during this specific timeframe. This table offers valuable insights into the model's robustness during the pre-Covid era and provides a basis for comparison with other periods.

	Model 8.A	Model 8.B	Model 8.C	Model 8.D	Model 8.E	Model 8.F
Dep. Variable	Investment Activities	Investment Activities	Investment Activities	Investment Activities	Investment Activities	Investment Activities
	FEARS30	FEARS25	FEARS20	FEARS15	FEARS10	FEARS5
FEARS	0.000273*** (0.000102)	0.000223** (9.49e-05)	0.000320*** (0.000101)	0.000266*** (9.02e-05)	0.000273*** (9.39e-05)	0.000391*** (8.81e-05)
FCF	0.000547 (0.000972)	0.000544 (0.000972)	0.000549 (0.000973)	0.000543 (0.000973)	0.000544 (0.000973)	0.000556 (0.000973)
DR	-0.00378*** (0.00124)	-0.00378*** (0.00124)	-0.00377*** (0.00124)	-0.00378*** (0.00124)	-0.00378*** (0.00124)	-0.00373*** (0.00123)
Size	-0.00339*** (0.000432)	-0.00339*** (0.000433)	-0.00338*** (0.000433)	-0.00339*** (0.000433)	-0.00339*** (0.000433)	-0.00338*** (0.000433)
ROA	0.00185* (0.00102)	0.00185* (0.00102)	0.00185* (0.00102)	0.00185* (0.00102)	0.00185* (0.00102)	0.00185* (0.00102)
MB	5.89e-05*** (1.63e-05)	5.89e-05*** (1.63e-05)	5.89e-05*** (1.63e-05)	5.89e-05*** (1.63e-05)	5.90e-05*** (1.63e-05)	5.90e-05*** (1.63e-05)
Age	-6.63e-05 (8.16e-05)	-6.76e-05 (8.16e-05)	-6.86e-05 (8.16e-05)	-6.70e-05 (8.16e-05)	-6.84e-05 (8.16e-05)	-6.86e-05 (8.15e-05)
BoardSize	0.000305** (0.000155)	0.000305** (0.000155)	0.000305** (0.000155)	0.000305** (0.000155)	0.000305** (0.000155)	0.000306** (0.000155)
Ind	-0.00394 (0.00250)	-0.00394 (0.00250)	-0.00393 (0.00250)	-0.00394 (0.00250)	-0.00394 (0.00250)	-0.00394 (0.00249)
Dual	0.00107 (0.000656)	0.00107 (0.000656)	0.00107 (0.000656)	0.00107 (0.000656)	0.00107 (0.000656)	0.00108 (0.000656)
Constant	0.0341*** (0.00399)	0.0342*** (0.00399)	0.0342*** (0.00399)	0.0342*** (0.00399)	0.0342*** (0.00399)	0.0341*** (0.00399)
Observations	17,784	17,784	17,784	17,784	17,784	17,784
R-squared	0.041	0.041	0.042	0.042	0.042	0.042
Standard Errors	Clustered (firm)	Clustered (firm)	Clustered (firm)	Clustered (firm)	Clustered (firm)	Clustered (firm)
Number of id	741	741	741	741	741	741

Table 14 - Robustness tests of Model 9

Table 14 demonstrates the impact of various FEARS index configurations on Model 9 during the Covid period. The table presents a range of potential scenarios under the pandemic conditions, highlighting the robustness and adaptability of Model 9. It provides valuable insights into the model's functionality and reliability amidst the unique circumstances posed by the Covid period.

	Model 9.A	Model 9.B	Model 9.C	Model 9.D	Model 9.E	Model 9.F
Dep. Variable	Investment Activities	Investment Activities	Investment Activities	Investment Activities	Investment Activities	Investment Activities
	FEARS30	FEARS25	FEARS20	FEARS15	FEARS10	FEARS5
FEARS	0.000520*** (0.000139)	0.000484*** (0.000125)	0.000604*** (0.000140)	0.000458*** (0.000105)	0.000416*** (9.38e-05)	0.000367*** (9.27e-05)
FCF	0.000998 (0.00172)	0.000986 (0.00172)	0.000989 (0.00172)	0.00100 (0.00172)	0.00100 (0.00172)	0.00102 (0.00172)
DR	-0.000442 (0.00214)	-0.000439 (0.00214)	-0.000323 (0.00215)	-0.000312 (0.00215)	-0.000312 (0.00215)	-0.000476 (0.00214)
Size	-0.00384*** (0.000997)	-0.00384*** (0.000997)	-0.00388*** (0.000996)	-0.00388*** (0.000996)	-0.00386*** (0.000993)	-0.00379*** (0.000990)
ROA	-0.000339 (0.00133)	-0.000331 (0.00133)	-0.000306 (0.00133)	-0.000307 (0.00133)	-0.000302 (0.00133)	-0.000342 (0.00133)
MB	1.72e-05 (2.08e-05)	1.73e-05 (2.08e-05)	1.64e-05 (2.08e-05)	1.64e-05 (2.08e-05)	1.63e-05 (2.08e-05)	1.73e-05 (2.08e-05)
Age	0.00158*** (0.000302)	0.00160*** (0.000299)	0.00162*** (0.000293)	0.00161*** (0.000291)	0.00164*** (0.000294)	0.00140*** (0.000262)
BoardSize	-8.11e-05 (0.000248)	-8.12e-05 (0.000248)	-8.02e-05 (0.000248)	-8.01e-05 (0.000248)	-8.05e-05 (0.000248)	-8.22e-05 (0.000248)
Ind	-0.000812 (0.00338)	-0.000813 (0.00338)	-0.000829 (0.00338)	-0.000829 (0.00338)	-0.000823 (0.00338)	-0.000793 (0.00338)
Dual	-0.000211 (0.000561)	-0.000212 (0.000561)	-0.000212 (0.000561)	-0.000212 (0.000561)	-0.000213 (0.000562)	-0.000213 (0.000561)
Constant	-0.0330** (0.0129)	-0.0339*** (0.0128)	-0.0343*** (0.0126)	-0.0338*** (0.0126)	-0.0352*** (0.0128)	-0.0252** (0.0117)
Observations	5,928	5,928	5,928	5,928	5,928	5,928
R-squared	0.033	0.033	0.034	0.034	0.034	0.033
Standard Errors Clustered (firm)	Clustered (firm)	Clustered (firm)	Clustered (firm)	Clustered (firm)	Clustered (firm)	Clustered (firm)
Number of id	741	741	741	741	741	741

Appendices

Appendix 1 - Summary of the Empirical Review

Appendix 1 provides a comprehensive overview of various studies. It categorizes and summarizes the key details of each study, including the sample size, the period of study, the market of interest, and the specific sentiment measure used. This table serves as a concise reference to understand the methodology and scope of each study under review.

Study	Sample, Period & Market	Sentiment Measure
Baker & Wurgler.(2007). Investor sentiment in the stock market	n.a	Employ six proxies (e.g Trading volume & IPOs first-day returns): BW Sentiment Index
Tetlock (2007). Giving content to investor sentiment: The role of media in the stock market.	Daily historical returns on Dow Jones Industrial Averages on U.S markets from 1984 - 1999, encompassing approximately 4,000 observations of market returns	Utilizes the GI to construct measure of media pessimism derived from the coverage in the Wall Street Journal's ("WSJ")
Tetlock et al. (2008). More than words: Quantifying language to measure firms' fundamentals	Over 350,000 qualifying news stories, 260,000 from DJNS & 90,000 from WSJ --> containing approx. 1 mn words. Firms listed on the S&P 500 from 1980 - 2004	Utilize a quantitative measure/textual analysis using the GI
Polk & Sapienza. (2009). The stock market and corporate investment: A test of catering theory.	Utilizes a sample composed by firms across the span of 1963 - 2000.	Tobin's Q
Li & Grundy. (2010). Investor sentiment, executive compensation, and corporate investment.	The sample comprises 20,289 firms spanning from 1992 - 2005.	Firm's Q ratio, representing the discrepancy between the market value and the replacement cost of the assets.
Da et al. (2015). The Sum of all FEARS.	Drawing upon the aggregated search volume of American households from 2004 - 2011	The FEARS Index
Zhu et al. (2018). Top management team demography, investor sentiment, and the investment levels of listed companies.	The sample is compromised by A-share companies listed on the Shanghai and Shenzhen markets from 2008 - 2012.	Investor sentiment is measured as a momentum index based on stock returns
Du & Hu. (2019). Why does stock-market investor sentiment influence corporate investment?	Analyze a sample of U.S. firms during the period of 1967 - 2015	Similar to Baker & Wurgler (2007): BW Sentiment Index
Li & Grundy. (2023). The effect of investor sentiment and the structure of shareholder ownership on corporate investment.	The sample encompasses US data spanning from 1980 to 2018.	Similar to Baker & Wurgler (2007): BW Sentiment Index

Appendix 2 - The North American Industry Classification (NAICS) codes and amount of study sample firms

Appendix 2 provides a detailed breakdown of the study's sample firms according to their respective industry classifications. The industries are organized by their 2-digit NAICS codes, providing a standard representation of economic sectors in North America. Alongside each industry and its associated code, the number of firms from the study sample that belong to that industry is listed. This appendix serves as a valuable reference for understanding the industry distribution and composition of the study sample.

Industry Classification	Codes (2-digits)	Number of firms
Agriculture, Forestry, Fishing & Hunting	11	8
Mining	21	15
Construction	23	11
Manufacturing	31-33	428
Wholesale Trade	42	20
Retail Trade	44-45	45
Information	51	66
Professional, Scientific & Technical Services	54	80
Administrative & Support & Waste Management & Remediation Services	56	14
Educational Services	61	5
Health Care & Social Assistance	62	8
Arts, Entertainment & Recreation	71	6
Accommodation & Food Services	72	21
Other Services (except Public Administration)	81	14

Source: U.S. Census Bureau. (n.d.). North American Industry Classification System (NAICS).

Appendix 3 - Variable definition table

Appendix 2 provides the definition for each of the variables used in our analysis. The first column provides the full name of the variable and the second column presents the abbreviated name used in STATA. The third column describes how each variable is measured. This appendix delivers a comprehensive guide for understanding the variables employed in our research.

Variable	Symbol	Definition
Dependent variable		
<i>Investment Activities</i>	IA	Average investments next three quarters divided by total assets
Explanatory variable		
<i>Financial and Economic Attitudes Revealed by Search</i>	FEARS	The change in adjusted search volume for the thirty terms with the strongest relationship to market returns
Control variables		
<i>Free Cash Flow</i>	FCF	Twelve trailing months cash flow from operations divided by total assets the quarter prior to the twelve trailing month period
<i>Debt Ratio</i>	DR	Total debt divided by total assets
<i>Firm Size</i>	Size	Natural logarithm of total assets
<i>Return on Assets</i>	ROA	Net income divided by total assets
<i>Market-to-Book ratio</i>	MB	Market value of equity divided by the book value of equity
<i>Firm Age</i>	Age	Years since foundation as of start of the year
<i>Board Size</i>	BoardSize	Number of directors serving on the board
<i>Board Independence</i>	Ind	Proportion of independent directors serving on the board
<i>Chair Duality</i>	Dual	Dummy variable that equals 1 if CEO and chair is the same person and zero otherwise

Appendix 4 - Summary statistics by size

Appendix 4 presents a complementary summary statistics by size for **Table 1**. The complementary summary statistics are categorized into two groups according to size, using a dummy variable assigned a value of 1 (0) when a firm's total assets exceed (fall below) the median total assets. All accounting-related variables (FCF, DR, TA, ROA, MB) have been winsorized at the top and bottom 1%. As predicted, the small firms have average total assets of \$171m, whereas the large firms' equivalent is \$6,956.67m. Larger firms typically carry more debt, which aligns with the notion that assets in place possess greater debt capacity than growth assets. Notably, more than half of the sample firms generate net negative returns on their capital and significantly lower cash flows than their larger counterparts. Small and large firms also differ in governance practices. On average, larger firms tend to have larger boards and higher board independence.

Variable	Symbol	Mean	Median	SD	Min	Max	N
Sizedummy: 0							
Dependent variable							
Investment Activities	IA	.01	0	0.01	0	.06	11841
Main Explanatory Variable							
Financial and Economic Attitudes Revealed by Search	FEARS30	-.03	-.04	0.39	-1.18	1	11872
Control Variables							
Free Cash Flow	FCF	-.1	.02	0.36	-1.4	.43	11868
Debt Ratio	DR	.17	.07	0.25	0	1.25	11872
Total Assets	TA	171	134.13	134.10	7.23	486.57	11872
Return on Assets	ROA	-.18	-.03	0.38	-1.49	.37	11841
Market-to-Book ratio	MB	3.49	2	7.71	-22.65	46.96	11872
Firm Age	Age	35.67	28	25.76	0	169	11872
Board Size	BoardSize	6.94	7	1.58	2	13	11872
Board Independence	Ind	.80	.83	0.10	0	1	11872
Chair Duality	Dual	.34	0	0.47	0	1	11872
Sizedummy: 1							
Dependent variable							
Investment Activities	IA	.01	.01	0.01	0	.06	11871
Main Explanatory Variable							
Financial and Economic Attitudes Revealed by Search	FEARS30	-.05	-.06	0.41	-1.18	1	11872
Control Variables							
Free Cash Flow	FCF	.11	.11	0.12	-1.4	.43	11871
Debt Ratio	DR	.27	.26	0.22	0	1.25	11871
Total Assets	TA	6956.67	1846.97	14570.05	486.73	77093	11871
Return on Assets	ROA	.05	.05	0.11	-1.12	.37	11871
Market-to-Book	MB	4.69	3.02	7.44	-22.65	46.96	11871
Firm Age	Age	43.88	36	29.75	0	171	11872
Board Size	BoardSize	8.95	9	1.96	4	18	11872
Board Independence	Ind	.84	.88	0.08	.43	1	11872
Chair Duality	Dual	.34	0	0.47	0	1	11872

Appendix 5 - Top thirty terms in absolute values for the FEARS index

Appendix 5 presents the top thirty terms with the highest absolute t-values in the FEARS index in the first quarter of 2014 (Q1-14). A larger absolute value of the t-statistics signifies a stronger relationship between the corresponding search term and CCMP market returns. This appendix allows readers to understand which search terms greatly influence the FEARS Index, providing a clearer picture of the factors affecting this measure.

Terms (1-15)	T-value	Terms (16-30)	T-value
<i>Liquidation</i>	0,63	<i>File bankruptcy</i>	-0,72
<i>Social security benefit</i>	0,69	<i>Bankruptcy court</i>	-0,58
<i>Bank bonus</i>	-0,47	<i>Bankruptcy chapter7</i>	0,89
<i>Charitable donations</i>	0,39	<i>Profit margin</i>	-0,32
<i>Charitable deduction</i>	-0,60	<i>Fraud</i>	-0,22
<i>HSA</i>	0,06	<i>Home equity</i>	0,14
<i>Scarcity</i>	0,57	<i>Procter & gamble stock</i>	-0,27
<i>Silver</i>	-1,01	<i>Jobless claims</i>	0,22
<i>LLC</i>	-0,06	<i>Jobless rate</i>	-0,05
<i>Nonprofit</i>	-0,55	<i>Poverty level</i>	0,20
<i>Profitable business</i>	0,13	<i>Federal poverty level</i>	-0,56
<i>Savings account interest</i>	-0,65	<i>Economic recession</i>	-0,23
<i>Insurance subsidy</i>	0,60	<i>Income</i>	0,54
<i>Dollar</i>	-0,34	<i>Health insurance</i>	0,30
<i>Bankruptcy</i>	-0,48	<i>Insurance for the unemployed</i>	0,03

Note: The terms are not ranked in chronological order.