DOCTOR OF PHILOSOPHY

# The Influence of Investor Sentiment on Mutual Fund Portfolio Composition

Lu, Qixin

*Award date:*
2025



*Awarding institution:*
University of Dundee

# The Influence of Investor Sentiment on Mutual Fund Portfolio Composition

Qixin Lu

Supervised by Dr. Martin Jones and Dr. Andrzej Kwiatkowski

Submitted to

School of Business

at

University of Dundee

In partial fulfilment of the requirements for the degree of

Doctor of Philosophy

2025

# Acknowledgements

I am honoured to learn and work with my supervisors, Dr. Martin Jones and Dr. Andrzej Kwiatkowski. I am indebted to Dr. Martin Jones and Dr. Andrzej Kwiatkowski for their support and their comments and feedback on my thesis. I have benefited from their incisive criticism and comments on my English. I would like to thank all the seminar participants at the University of Dundee. I would like to thank Prof. Sudharshan Reddy Paramati for his encouragement and advice. Additionally, I am incredibly grateful for my parents for their continuous support.

# Declaration

I hereby declare that the work in this thesis was carried out in accordance with the Regulations of the University of Dundee. I have not used any published materials in this thesis.

Three empirical chapters are single authored:

“The Relationship Between Multiple-Based Investor Sentiment Proxies and Mutual Fund Portfolio Composition” (Chapter 3)

“Will a Portfolio of Mutual Funds React to Investor Sentiment Fluctuations Asymmetrically?” (Chapter 4)

“A New Composite Investor Sentiment Index: Does It Affect Mutual Funds’ Portfolio Composition?” (Chapter 5)

# Abstract

This thesis addresses the relationship between investor sentiment and mutual fund portfolio composition and investigates whether prospect theory explains the behaviours of institutional investors like mutual funds. Most of the existing literature focuses on mutual funds' performance related to investor sentiment or uses a mutual fund-based sentiment proxy. There is limited research on how investor sentiment impacts the composition of mutual fund portfolios. To fill this gap, this study creates a new composite investor sentiment index that includes survey-based, market-based, and internet-based sentiment indices and compares the effectiveness of the various sentiment measurement methods in reflecting mutual fund portfolio composition. This focus aligns with the broader objective of elucidating the behavioural underpinnings of investment decisions in the face of market uncertainties.

In the first empirical chapter, I examine the different effect of multiple investor sentiment proxies on mutual fund portfolios based on the panel ARDL method. The portfolio compositions are represented by two dependent variables: the percentage of capital invested in the common stock market and the percentage of cash holdings in mutual funds, respectively. The results indicate that all investor sentiment indicators will have an influence on mutual fund portfolio composition and their investment decisions; however, some of them have a more profound effect during crisis periods than stable periods. In addition, the results suggest that the composite sentiment index outperforms single indices when impacting mutual funds' portfolios, and the survey-based index will significantly influence mutual funds' investment decisions in both crisis and stable periods. Nevertheless, the internet-based and market-based indices have a partial effect when the market is in crisis or remains stable, but they still carry

information about how they significantly influence mutual fund portfolio composition.

The second empirical chapter applies the nonlinear panel ARDL method to investigate asymmetric effects of investor sentiment on mutual fund portfolio composition. In addition, I investigate whether the prospect theory explains the behaviours of institutional investors. The empirical results imply that mutual funds may have different reactions to increases and decreases in investor sentiment. The observed reactions can be rationalized using the prospect theory, which proposes that people may have different subjective valuations towards gains and losses. This indicates that mutual fund investors might value losses more than gains and suffer from loss aversion bias.

The last empirical chapter employs the principal component analysis (PCA) method to build a composite investor sentiment index, consisting of various sentiment proxies such as surveys, the Internet, and market-based measures. The proposed index may carry more information than the well-known BW Index, constructed based on pure market-based sentiment indices. Furthermore, this chapter also applies panel ARDL and nonlinear panel ARDL methods to investigate the influence of the new PCA composite index on mutual fund portfolio composition. The results indicate that the new sentiment index has a significant correlation with both categories of mutual fund portfolio allocation and has a strong relationship than the BW Index. The results of the nonlinear analysis also indicate that mutual funds react to this index asymmetrically and consistently with the value function proposed by prospect theory.

The collective results suggest that all investor sentiment indicators can significantly impact mutual fund portfolio composition. Shifts in investor sentiment, whether positive or negative, have disparate effects on mutual funds' allocations of assets between stocks and cash. These findings may have important implications for investing decisions and portfolio management strategies.

# Contents

# Chapter 1. Introduction

Recent advances in behavioural economics have shown that 'investor sentiment' can affect many aspects of the economy or financial sector. Behavioural biases are essential to consider when understanding anomalies in economic activities. Contemporary research reveals that irrational investor sentiment may influence institutional investors' decision-making and may have contributed to one of the greatest economic crises of the 21st century, during 2007-2008(Bekiros et al., 2017). The irrational sentiment derived from herding behaviour can lead to a market bubble and prevent individuals from generating rational estimations, hence fuelling the effect of market risk. Therefore, it is possible that herding behaviour, driven by investor sentiment, played a role in triggering the 2008 crisis.

In addition, Baker and Wurgler (2006)find that sentiment generates cross-sectional variances in stock returns. Their research suggests that stock returns may be negatively related to market sentiment levels and that low-sentiment beta stocks may outperform high-sentiment beta stocks. The reason for this phenomenon is that low-sentiment beta stocks are usually underestimated, and their subsequent returns are higher than those of overvalued stocks with a positive loading on sentiment. Furthermore, DeVault et al. (2019) and Wang (2020) show that though institutional investors are professional, they are also prone to investor sentiment when facing market fluctuations. This feature is observable in fund managers designing better portfolio strategies to maximize profits, for example, not being misled by sentiment and making biased decisions but rather exploiting sentiment in others to earn a profit.

Numerous scholarly works detail the correlation between sentiment and mutual funds. Mutual fund flow can be used as an indicator for monitoring changes in investor sentiment between bond funds and equity funds. Expressly, changes in mutual fund flow can reveal whether investors are feeling optimistic or pessimistic about the stock market(Ben-Rephael et al., 2012). Moreover, Nguyen et al. (2018)find that mutual fund performance relates positively to investor sentiment, which confirms that behavioural factors play a significant role in asset return. However, most of the literature focuses on mutual funds' performance related to investor sentiment or uses a mutual fund-based sentiment proxy. There is limited research on how investor sentiment impacts the composition of mutual fund portfolios.

The research motivation and research gaps are as follows. Most of the existing literature on investor sentiment is relevant to investigating sentiment's influence on stock market returns; however, there are several gaps in the literature, such as the impact of investor sentiment on portfolio composition. The relationship between investor sentiment and portfolio composition can be evaluated from both the individual investors and institutional investor's perspectives. However, it is difficult and unrealistic to access specific personal portfolio construction data; therefore, the mutual fund portfolio composition data are employed in this study as substitutes. Mutual fund managers may respond to investor sentiment by adjusting their portfolios to cater to the preferences of their investors. For example, fund managers may increase their exposure to higher-risk assets to meet the demand of fund investors(Massa and Yadav, 2015).

Concerning the relationship between mutual funds and investor sentiment, the existing literature primarily focuses on the nexus between investor sentiment and portfolio return or regards mutual fund flows as a sentiment indicator. For example, Ben-Rephael et al. (2012)discover that mutual fund flow can be used to gauge stock market sentiment in the US. According to Indro (2010) empirical study, net aggregate equity fund flows are related to the sentiment of both individual investors and newsletter writers. Additionally, the research of Nguyen et al. (2018), they focus on the connection between investor sentiment and mutual funds' performance in the market of India and Pakistan, and detect the positive relationship between investor sentiment and fund performance. DeVault et al. (2019)state that sentiment metrics, which have traditionally been assumed to capture the behaviour of individual (irrational) investors, actually capture the demand shocks on stocks of institutional investors. This challenges the conventional wisdom that individual investors are the primary drivers of sentiment-based demand shocks. However, there has been little research on how the composition of mutual fund portfolios, such as the proportion of capital invested in different asset categories, changes as sentiment fluctuates. Moreover, few studies have concentrated on the connection between investor sentiment and existing portfolios in the market, such as individual or institutional portfolios like mutual funds.

Some literature focuses on the relationship between portfolio performance and sentiment; however, most of the portfolios in the research are designed or hypothesized simulations. For instance, Frugier (2016)selects 46 stocks from the European stock market to compose several portfolios to investigate whether returns differ when considering investor sentiment, rather than using a mutual fund portfolio operating in the market. Hence, Frugier (2016)research only concludes how sentiment might impact a portfolio at a fixed time point, whereas my study provides a long-term vision of dynamic influence.

Bu and Forrest (2021)and Bu (2021) only compare a limited number of sentiment measures reflecting mutual fund performance rather than portfolio composition. For instance, the direct measure like survey of American Association of Individual Investor (AAII) may outperform the indirect sentiment measure, the BW Index, in predicting stock and mutual fund performance.

Additionally, Chue and Mian (2022)only present two sentiment indicators. They consider how the indicators affect mutual fund portfolio activeness by using 'active share', the percentage of a fund's portfolio that differs from its benchmark index, as a dependent variable, which reflects the effort of fund managers to beat the market. They do not investigate the various basis indexes or determine which can most accurately reflect the portfolio composition. Furthermore, because their sample period ranges from 1985 Q1 to 2009 Q3, the timeliness of their research is limited. In contrast, this study considers an internet-based sentiment measurement and employs different direct sentiment measures and multiple indirect sentiment measures to compare their correlation toward mutual fund portfolio composition.

This study contributes to the literature in a few ways. Although Cornell et al. (2011) show that sentiment impacts institutional investors' stock holdings, this study employs two dependent variables representing two different categories of holdings, providing a more specific inside view of how institutional investors alter their portfolio, such as by transferring funds from high-risk assets to low-risk assets during sentiment fluctuation. If the percentage of stocks increases, this suggests that mutual funds are risk-seeking. In contrast, if the cash holdings increase, this indicates that mutual funds are risk averse. Additionally, this study compares different-based sentiment indicators and demonstrates the effectiveness of different-based sentiment indexes in reflecting mutual fund portfolio composition. This study also can confirm whether sentiment indicators are a beacon or a signal and offers a direction for less-informed individual investors. Individual investors possess low-level financial literacy and knowledge to modify their portfolio composition during a period of rapid sentiment fluctuation. Mutual funds are usually managed by highly sophisticated and well-informed professional teams, and their approach is designed to avoid severe losses.

Second, this study sheds light on whether investor sentiment affect the mutual fund portfolio composition asymmetrically and detects whether institutional investors are impacted by the cognitive bias loss aversion, demonstrated in the value function proposed by prospect theory. This study reinforces the power of sentiment indicators and the application of prospect theory in portfolio construction.

Finally, this study creates a new composite investor sentiment index that includes multiple-based sentiment indices. This may include survey-based, internet-based, and market-based measurements that gather more information than solely market-based indicators such as the BW Index.

Hence, this thesis tests following hypothesis:

Hypothesis 1: Different investor sentiment proxies will significantly affect mutual fund portfolio composition.

Investor sentiment, often quantified through various proxies, is known to influence market behaviours and decisions. These proxies, which include indicators such as investor surveys, market indices and Internet-based index, can have differential effects on portfolio choices made by mutual fund managers. According to studies like those by Brown and Cliff (2005) and Baker and Wurgler (2006), different investor sentiment proxies are linked to changes in risk-taking behaviour, which, in turn, could lead to different portfolio compositions in mutual funds. Mutual fund managers might adjust their portfolio allocations in response to these sentiment measures, either by shifting towards safer assets in times of low sentiment or increasing risk exposure when sentiment is high.

Hypothesis 2: Mutual fund portfolio composition will react asymmetrically to investor sentiment measures, consistent with the loss aversion principle of prospect theory.

Prospect theory, developed by Kahneman and Tversky (1979), suggests that investors are more sensitive to potential losses than equivalent gains—known as loss aversion. This implies that mutual fund managers may adjust their portfolio compositions more aggressively in response to negative sentiment than to positive sentiment. The asymmetry in reaction could result in a more pronounced shift towards conservative investments when sentiment declines, compared to a relatively modest increase in risk exposure when sentiment improves. This hypothesis aligns with findings from Barberis et al. (1998), who argue that investor psychology, particularly loss aversion, plays a crucial role in shaping market outcomes and investment strategies.

Hypothesis 3: The newly constructed PCA sentiment index (via the Baker and Wurgler method) will impact mutual fund portfolio composition with an asymmetric effect.

Principal Component Analysis (PCA) has been utilized in finance to combine multiple sentiment indicators into a single, composite index, as demonstrated in the methodology by Baker and Wurgler (2006). This composite sentiment index aims to capture the multiple based sentiment proxies like survey-based market-based and Internet-based, making it a powerful tool for reflecting investment behaviour. The hypothesis posits that this PCA-based sentiment index will influence mutual fund portfolio composition, and mutual funds portfolio composition will also react to it asymmetrically.

In this thesis, I comprehensively address investor sentiment's influence on mutual fund portfolio composition in the US market. There are many sentiment metrics, such as single indicators or composite indicators, but which of them best reflects the underlying sentiment information? Accordingly, the first empirical chapter, Chapter 3, investigates how different investor sentiment indicators influence mutual fund portfolio allocation. Investor sentiment needs time to reveal its effect on the market, which is why I applied the panel autoregressive distributed lag (panel ARDL) methodology to examine the power of investor sentiment toward mutual fund portfolios. I employed multiple investor sentiment measurements, including the consumer confidence index conducted by a non-profit, nonpartisan organization that provides research and insights on business, economics, the US Conference Board, the Baker and Wurgler (BW) Index, the Google Trends search volume of 'financial crisis' and 'economic recession', the number and first-day returns of initial public offerings (IPOs), and the Chicago Board Options Exchange (CBOE) volatility index, to estimate their strength in reflecting changes in mutual fund portfolio composition.

Investor sentiment measures a behavioural factor; nevertheless, there may be an asymmetric effect relating to gains and losses represented by investor sentiment, which might impact investment decisions. For example, the prospect theory proposed by Kahneman and Tversky (1979) implies that people may react to gains and losses differently due to their different subjective valuations. Thus, in the second empirical chapter, Chapter 4, I examine the investor sentiment proxies employed in Chapter 3 using a nonlinear panel ARDL methodology, to detect whether there is an asymmetric effect from investor sentiment to mutual fund portfolio composition. In addition, I investigate whether prospect theory can explain this circumstance, whether the behaviours of professional institutional investors like mutual funds are consistent with prospect theory.

Moreover, Baker and Wurgler (2006)present a sentiment measure based on the first principal component of standardized sentiment proxies employing six macroeconomic indicators, which include the closed-end fund discount, New York Stock Exchange (NYSE) share turnover, the number of IPOs, the average first-day returns, the share of equity issues in total equity and debt issues, and the dividend premium. However, investor sentiment is volatile, containing both rational and irrational components; as a result, it is difficult for purely market-based measurements to fully capture investor sentiment. Hence, to address this issue, I built a new composite index using a principal component analysis (PCA) methodology in the last empirical chapter, Chapter 5. This index includes five investor sentiment proxies with different bases, such as the consumer confidence index, the consumer sentiment index, Google Trends search volume, the number of IPOs, and the CBOE volatility index. After constructing the index, I applied the methodology used in Chapters 3 and 4 to estimate the strength of the new index and reveal mutual fund portfolio composition. In addition, Chapter 5 identifies whether the new PCA index has a nonlinear relationship with portfolio composition or fits into prospect theory. Overall, the thesis is organized as follows:

Chapter 2 is the literature review, which introduces the development of behavioural economics and elaborates on the definition of investor sentiment and different measurements of investor sentiment. In addition, Chapter 2 introduces the existing literature on portfolio composition, especially mutual fund portfolios, and linking prospect theory to the relationship between investor sentiment and mutual fund portfolios. Chapter 3 examines the impact of multiple-based sentiment indicators, like survey-based, market-based, or internet-based sentiment indicators, on mutual fund portfolios. Chapter 4 examines the nonlinear effects or fluctuations of different investor sentiments on mutual fund portfolio composition. Chapter 5 constructs a new sentiment index including survey-based, market-based, and internet-based sentiment indicators and examines the power of this new composite index to reflect mutual fund portfolio allocation both linearly and nonlinearly. All empirical chapters are accompanied by their applied methodologies. Chapter 6 is the conclusion, which describes the main findings of this thesis and the direction for future research.

# Chapter 2. Literature Review

## 2.1 Background

Portfolio construction has evolved and has been widely studied. The most well-known researcher, Markowitz (1952), is regarded as the father of modern finance. His theory, market portfolio theory, proposes that investors can establish their portfolio and gain the maximum expected return under certain levels of risk, and the more risk an investor is willing to take, the higher the return they might achieve.

After market portfolio theory, two other widely-known theories appear, the capital asset pricing model (CAPM) (Treynor, 1961; Sharpe, 1964; Mossin, 1966; Lintner, 1969)and the efficient market hypothesis(Fama, 1970). However, both fail to predict or explain the anomalies in some scenarios. For instance, it is challenging to eliminate overpricing, and a firm's stock price may reflect the views of investors who are too optimistic(Stambaugh et al., 2012). These models of traditional theory can predict some behaviours, but they only have limited power to forecast the behaviours of professional economists. What about the people who comprise a large percentage of the market but are not professional economists? For example, Black (1986)postulates that the noise traders, who have limited ability to identify useful information and make decisions according to noise information with a lack of professional support or advanced fundamental analysis, trade impulsively in the market. Nevertheless, the existence of noise traders violates the above theories' assumptions that the information is publicly available, and all market participants are rational. Therefore, it is necessary to consider people-related factors when investigating financial areas.

## 2.2 Theoretical Framework

Behavioural economics and finance are attracting scholarly attention. It explores the psychological factors influencing investors' financial decisions, challenging the Efficient Market Hypothesis. A key concept within this domain is investor sentiment, which refers to the overall attitude of investors towards market conditions. Investor sentiment can drive market prices away from their intrinsic values, leading to phenomena such as bubbles and crashes. Shiller (2003)bridges the gap between efficient market theory and behavioural finance, arguing that investor sentiment plays a crucial role in market dynamics. His analysis illustrates the importance of psychological factors in financial markets, showing how collective sentiment can lead to significant mispricing. Baker and Wurgler (2007)focus on how investor sentiment influences the stock market. They argue that sentiment-driven behaviours can lead to varying risk-taking activities among investors, affecting portfolio composition. Ben-Rephael et al. (2012)empirically examines the link between mutual fund flows and investor sentiment, providing evidence that sentiment is a significant determinant of portfolio decisions. Their works provide a foundation for understanding how different sentiment proxies—such as surveys, market indices, and Internet-based indices—can influence mutual fund managers' decisions and support the hypothesis that different sentiment proxies can have differential effects on mutual fund portfolio composition, as fund managers adjust their strategies based on sentiment measures.

Next, Kahneman and Tversky (1979), who built the prospect theory in the 1970s, suggest that investors prefer value gains to losses and investigate the common ground between economic and psychological areas. The prospect theory provides a framework for understanding how people make decisions under uncertain circumstances, which include the value function and weighting function.

Prospect theory's value function describes how individuals value gains and losses. It is usually expressed in terms of changes (gains or losses) relative to a specific reference point, rather than final outcomes (Figure 2–1). The value function is typically concave for gains and convex for losses, reflecting diminishing sensitivity to changes in wealth. Moreover, it is steeper for losses than for gains of the same size, indicating that losses generally have more impact on individuals' feelings than equivalent gains; this is referred to as loss aversion.

Prospect theory's weighting function represents how people perceive the probabilities of potential outcomes when making decisions under uncertainty (Figure 2–2). Instead of treating probabilities in a linear or direct fashion as utility theory suggests, prospect theory suggests that people tend to overestimate the probability of unlikely events (thus giving them more 'weight') and underestimate the probability of likely events. The function is therefore not linear but involves a curve in which both low and high probabilities are over-emphasized and medium probabilities are under-emphasized. Barberis (2013) reviews the application of Prospect Theory in economics, emphasizing its relevance in explaining market anomalies and investor behaviours that deviate from traditional economic models. His assessment highlights how loss aversion can lead to asymmetric responses in financial markets, such as stronger reactions to negative news compared to positive news. Gu and Yoo (2021)extend the application of Prospect Theory to mutual fund flows, demonstrating that investors' loss aversion significantly influences the flow of funds into and out of mutual funds. The above findings provide the foundation of hypothesis that mutual funds, influenced by investor behaviour, may adjust their portfolio compositions more significantly in response to negative sentiment, aligning with the loss aversion principle.

At last, Baker and Wurgler (2006)demonstrate how PCA can be utilized to construct a composite sentiment index that integrates multiple sentiment proxies. This index is particularly powerful in reflecting investment behaviour, as it captures the broader sentiment that drives market dynamics. Principal Component Analysis (PCA) is a statistical technique used to reduce the dimensionality of data by combining various indicators into a single index. In the context of investor sentiment, PCA can synthesize multiple sentiment measures into a composite index, providing a comprehensive view of overall market sentiment. However, their index only includes market-based sentiment proxies. The hypothesis that the new PCA sentiment index which added survey-based and Internet-based sentiment indicators will impact mutual fund portfolio composition, with an asymmetric effect, is grounded in the understanding that this composite index offers a more nuanced measure of market sentiment compared to the Baker and Wurgler index. Fund managers can use this index to adjust their portfolios, anticipating that the market may react more strongly to negative sentiment than to positive sentiment, in line with loss aversion.

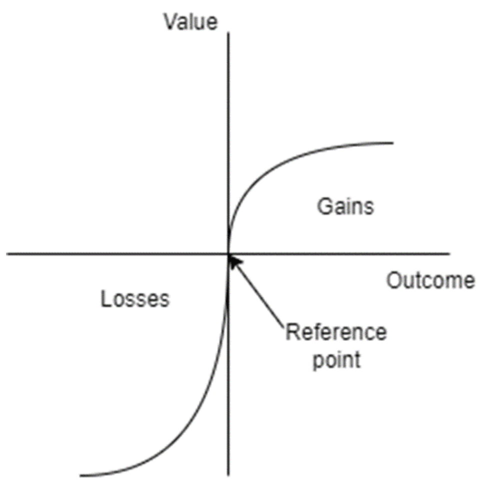



**Figure 2–1. Value function**

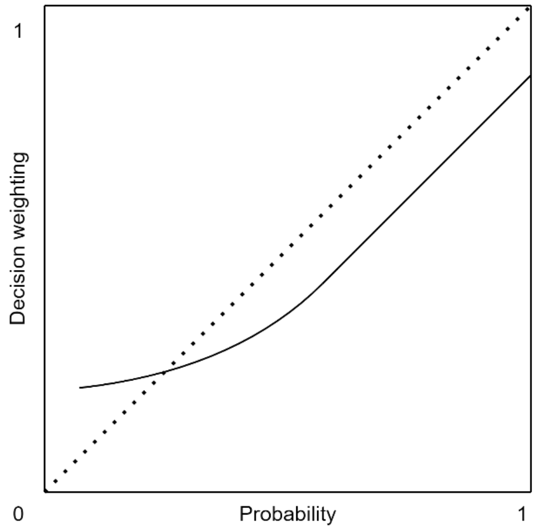



**Figure 2–2. Weighting function**

## 2.2 Understanding Behavioural Factors in the Market

After these developments in the field of behavioural finance, investor sentiment emerged to measure the anomalies detected in the behavioural finance area. Scholars have conducted many studies to find the effect of investor sentiment. Lee et al. (1991)research the relationship between investor sentiment and closed-end funds; sentiment may have a significant influence on closed-end funds. Fisher and Statman (2000)and Baker and Wurgler (2006) study the connection among investor sentiment and stock return, build a composite sentiment measurement, and detect a significant relationship with stock market return. Hence, investor sentiment is a significant factor impacting the market.

This chapter elaborates on the role of investor sentiment by discussing its background, such as how and why sentiment can generate significant influence on economic and financial performance, how to detect it in the market since its influence is hidden, or how it can be discovered systematically. However, behavioural finance theories are not accepted or given priority on this topic.

Nevertheless, behavioural finance tackles problems that traditional finance theories find difficult to explain, for example, the ‘noise traders’ detected by Kyle (1985) and Black (1986) and the excess volatility in stock market anomalies(Shiller, 1990).

Hence, the first section of this chapter provides an overview of fundamental and conventional financial theories along with their deficiencies, for instance, the market portfolio theory(Markowitz, 1952), the CAPM, and the efficiency market theory. It is not difficult to see that these theories need strong and unrealistic assumptions to stand, which is why behavioural finance may contradict traditional financial theories. The crucial issue concerns human factors, such as how investors act and think during market and economic activity. Does investor behaviour affect the final outcomes? This is for behavioural economics and finance theories to answer.

As far as this question is concerned, the second part of this chapter elaborates on the history and development of behavioural finance theories. It is vital to understand how a tool works and is generated. The behavioural economic theories, or to be more precise, some points of the theories, seem to have appeared in history far earlier, as the theories only appeared in mainstream discussions in recent decades. Smith states some interesting and important thoughts about economics concerning behavioural directions like overconfidence and loss aversion. Thaler (2016)believes the root of behavioural economics is in the work of Adam Smith.

Adam Smith roughly describes three key aspects that will develop in future behavioural economics: overconfidence, loss aversion, and self-control. On the concept of overconfidence, Smith (2010b)states that the excessive self-confidence that most individuals possess in their own abilities is a long-standing issue. However, their irrational confidence in their likelihood of success has received less attention.

As for the definition of loss aversion, Smith (2010a)does not describe loss directly, but uses fear of pain which may link to loss; he states that the aversion to pain stems from self-interest, and people acknowledge that their empathy toward it is not based on any anticipated benefits it might yield. Moreover, Smith mentions the self-control of men. 'Self-command is not only itself a great virtue, but from it all the other virtues seem to derive their principal lust' (Smith, 2010a)

Furthermore, Keynes (1937)captures the influences of sentiment and expectation have on economic activities, in which he discusses the role of 'animal spirits. It refers to the instinctual, emotional, and mental force behind human activity, particularly in economic decision-making. Animal spirits embody those non-rational motivational and confidence factors of an individual, especially investors, that influence their willingness to take risks, make investments, and engage in economic activities. This conception emphasizes the importance of psychological and emotional factors in economic activity, highlighting the limitations of purely rational models in explaining and predicting market behaviour.

The behavioural theories were stifled by mainstream economics for decades. Despite several voices supporting them during this period, they attracted limited attention. In the 1970s, Kahneman and Tversky (1979)start to combine psychology and economics for research, which inspired many scholars to step into the field of behavioural economics.

Behavioural economics and finance now propose that human behaviours impact economics as a cornerstone, and academia accepts this point. Therefore, the next significant step is to figure out what determines and generates behaviours. It is easy to understand that almost all behaviours and activities are generated by our thoughts or choices, like deciding to drink coffee during a work break, or having a cupcake after dinner. Individuals face choices and make different decisions all the time. In addition, they behave differently according to their options. The different behaviours may lead to different effects, like having trouble sleeping after that coffee or gaining weight after that cupcake.

This also works when applied to economic or financial behaviour. Therefore, it is necessary to discuss what contributes to the decision-making process. Decisions are impacted by inside and outside factors. Inside factors include beliefs, emotion acknowledgement, or anticipation, which belong to the psychological area, whereas outside factors include the environment or changing situations, like the fluctuation of an unstable financial market. For instance, the choices of mutual fund managers may be impacted by investor sentiment coming from different directions, like the market or internet. Combining multiple directions together, as a power, influences the economic and financial markets.

However, investor sentiment does not have a solid definition; there are different versions. At the heart of these discussions, two distinct elements continuously emerge: desire and belief. The desire aspect of investor sentiment is characterized by the anticipation of favourable investment outcomes. When people are expecting good news, they have optimistic anticipation, which involves a desire for the anticipated good scenario. However, when people think what will happen, that is a belief. Expressly, optimistic expectations are investors' desires, whereas pessimistic expectations are not what they believe will happen. According to the research of Bilel and Mondher (2021), investor sentiment contains the desire of investors, which demand more dividends. Baker and Wurgler (2007)state that investor sentiment reflects desire or expectation concerning future investment opportunities.

Other researchers state that investor sentiment is defined as the expectation or belief toward future cashflow, whether it be bearish or bullish, which refers to the market price being below or above the fundamental value(Brown and Cliff, 2004; Chang et al., 2012; Beer et al., 2018). In addition, Baker and Wurgler (2006; 2007)define investor sentiment as the optimistic or pessimistic attitude toward a specific asset, such as the stock market, Liu (2015)agrees with this definition. However, sentiment mainly reveals the desires of investor about the market.

Interestingly, these definitions converge on the view that investor sentiment encapsulates a blend of both desires and beliefs. This blend influences the subjective valuation derived from rational fundamental analysis. For example, when investors harbour a bearish belief, their desire to avoid losses may lead to loss aversion, a key concept in the prospect theory built by Kahneman and Tversky (1979).

## 2.3 Shortcomings of Traditional Economic Theories

In the finance field, there are many basic portfolio selection theories contributed by well-known scholars, such as Tobin (1958), Sharpe (1964), Lintner (1969), and Markowitz and Todd (2000). Market portfolio theory, CAPM, and efficient market hypothesis are cornerstones of modern finance. Market portfolio theory is presented by Harry Markowitz in his 1952 article, ‘Portfolio Selection’. He suggests a way for investors to build their portfolios under a certain level of market risk to achieve the maximum expected return, emphasizing that higher returns are always accompanied by risk. This theory confirms that the concept of the ‘efficient frontier’, which is an aggregation of efficient portfolios that suggests every point on the line, refers to an optimal investment portfolio that gives the highest expected return with the lowest possible risk. A graph can be plotted with every possible securities combination, and the X axis of the graph refers to the risk of this portfolio, whereas the Y axis represents the expected return. Hence, this plot demonstrates the most optimal and profitable portfolios. The efficient frontier can be drawn as an upward sloping hyperbola with every efficient portfolio on the line. Those combinations below the curve (see Figure 2–5) are not efficient and putting money into them is not rational. The merit of modern portfolio theory is that it provides the ability for investors to construct their portfolios not only on the peculiarities of the securities but also the co-movement between different securities. Therefore, portfolios that consider the interactions between securities have a higher expected return than those that do not, but they both share the same level of risk(Elton and Gruber, 1997).

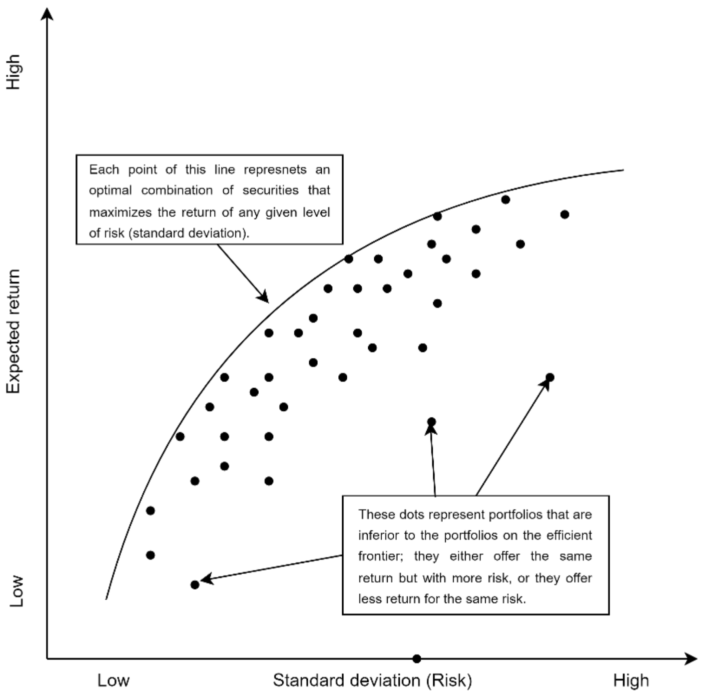



**Figure 2–3. The efficient frontier**

However, this theory has limitations; it only concerns selecting securities based on beliefs about the future performance of certain securities; it does not properly investigate and explain the root of those relevant beliefs according to observation. Therefore, Rice (2017)suggests that market portfolio theory omits many factors, such as geopolitical risks or duration, that push the market forward and are invisible to the calculation of market portfolio theory. In addition, he believes that applying the theory from 1952 to the present market is 'like going to a museum's map room to plan your next trip'. The 1952 theory overlooks critical changes in economics and markets like the digital economy or globalization. In addition, modern portfolio theory is a descriptive and theoretical attempt to illustrate the working process of the capital market rather than build an investment portfolio. According to Curtis (2004), the predictive power of modern portfolio theory about future market movements is quite low over any time horizon that considers human investors, and he believes the base assumptions that all investors are rational are clearly wrong because investors do not always act based on their rationality; rather, they make decisions based on biases, emotions, or other psychological factors that do not align with rational economic principles.

During the 1960s, Treynor (1961), Sharpe (1964), Mossin (1966) and Lintner (1969) created the capital asset pricing model (CAPM). This theory is widely used in the fields of investment decision-making and corporate finance, providing a vital method for market participants to form investment combinations. CAPM theory describes the link between risk and expected return of assets, especially stocks, provides a method for investors to calculate risk and how much profit they should expect, and decides in which securities they should invest. The equation of the CAPM is as follows:

$$E_{r,i} = r_f + \beta_{i,m}\left(E_{r,m} - r_f\right)$$

where $E_{r,i}$ represents the expected return rate of asset $i$; $r_f$ represents the risk-free return rate; $\beta_{i,m}$ represents the beta of asset I, and $\left(E_{r,m} - r_f\right)$ represents the market risk premium.

CAPM theory is based on two strong assumptions: the market is efficient, and investors are all rational and risk averse. However, this model may be too idealized, which causes researchers to question the model. Fama and French (2004)admit that the empirical failure of CAPM results from these assumptions. Hence, from a market structure aspect, there are several unrealistic assumptions, such as that all assets are publicly held and trade on public exchanges, and investors can borrow or lend at the same risk-free rate. In addition, information access is costless and publicly available, which means investors are fully informed and should have homogeneous expectations (i.e., the same estimates of expected returns, variances, and covariances), which is impossible in the real world. As for the market's ability to process enough liquidity without any taxes or transaction costs, the opinions of financial institutions should be considered. As far as individual behaviour is concerned, all investors under CAPM are supposed to be rational and mean-variance optimizers. Additionally, they chase the optimal expected utility of wealth over a single-period investment horizon. These individual assumptions prompt arguments among behavioural finance advocates.

It is easy to understand why CAPM fails to fit into a realistic market under these conditions. Guo (2004)aptly argues that the CAPM's failure in real-world markets can be attributed to its reliance on unrealistic assumptions, such as the assumption of fully rational individual investors. Moreover, an empirical examination of the CAPM by Basu and Chawla (2010) in the Indian market, encompassing 10 portfolios over a five-year timeframe, further corroborates the model's inability to demonstrate market efficiency and effectiveness. These findings highlight the limited explanatory capability of the CAPM in real-market contexts. Therefore, De Bondt and Thaler (1995)humorously joke that 'finance consists of theories for which there is no evidence and empirical facts for which there is no theory'. Moreover, in behavioural finance, scholars suggest that CAPM neglects the effect of individuals' behaviours. Shefrin and Statman (1994)propose that the failure of the CAPM is caused by the interaction of noise traders who trade with low-quality information and make epistemic mistakes. Ricciardi (2008)emphasizes that the failure of the CAPM might be the result of overlooking a qualitative aspect of behavioural risk, such as the influence of cognitive issues or highly significant emotional factors, which are covered by the assumptions of the CAPM.

After the CAPM, the efficiency market theory was introduced by Fama in 1970, which marked progress in modern financial theory. According to Fama (1970), the efficiency of securities prices in fully reflecting the information that exists in the market determines the 'efficiency' of the market. However, this hypothesis requires certain solid conditions to be met. First, when new information emerges, the market will immediately react and adjust to the new price. The movement of prices follows a random walk pattern, which is influenced by the occurrence of new information. Second, new information is inherently unpredictable, such as the sudden outbreak of the coronavirus. Is it good news or bad news? Time will tell. Last, investors always pursue the most profitable opportunities, and their analysis of securities is independent and not influenced by others.

The efficient market hypothesis comprises three forms: weak form efficiency, semi-strong form efficiency, and strong form efficiency. Weak form efficiency suggests that market prices reflect all historical information, including strike prices and trade volume. Semi-strong form efficiency indicates that market prices reflect all publicly available information, such as company profitability or disclosed financial information. However, it is important to note that information that is not publicly available can still lead to significant profits under this form. Strong form efficiency implies that market prices reflect all information about a company's operations, regardless of whether it is public or private. Nevertheless, it is crucial to recognize that the information on the market may not be as useful as individuals believe. Kyle (1985)and Black (1986) propose the concept of 'noise traders', those who lack the ability to access inside information and regard the 'noise' as information to strike their investment, which violates the assumption of Fama's model that information is publicly available to all market participants.

Nevertheless, the efficiency market hypothesis is not equipped with enough power of explanation for all the anomalies that exist in the market, such as excess volatility in the stock market. (Shiller, 1990)concludes that different investors may hold different beliefs about the different models they employ, and that will influence their behaviours and judgments concerning their expected return. Caparrelli et al. (2004)conduct research on the Italian stock market and detect herding activity, which disagrees with the efficient market hypothesis. Herding indicates that irrational investors follow the crowd to strike blindly, which suggests that the market is inefficient.

Moreover, one of the most fundamental assumptions of the efficient market hypothesis is that all market participants, including small household investors, speculators, and big fund managers, are rational when making decisions. However, it is obvious that it is not possible for individuals to maintain a consistently clear mindset to calculate or design every investment strategy, which would require all investors to be sophisticated economists. Ritter (2003)rightly states that even sophisticated economists are not completely rational because of their own preferences and mistaken beliefs. Personal preference does play a part in investors' beliefs, desires, and decision-making. (Golman et al., 2017)study demonstrates that individuals tend to avoid information that is contrary to their interests, even if it is both useful and freely available. Furthermore, individuals may also refrain from considering strategic factors due to the influence of hedonic motives. For instance, managers often avoid hearing arguments that contradict their initial decisions. Therefore, the findings indicate that preferences play a significant role. Rather than actively seeking useful information, people tend to avoid it, even when they are aware of its availability, opting instead for information that aligns with their own biases and desires.

The return rate of stocks may continue to move in its original direction; for instance, historical high-return stocks will still earn more profit than historical low-return stocks in the future, a phenomenon called the momentum effect detected by Jegadeesh and Titman (1993). The momentum effect is the rate at which the securities' price or volume are overpriced. It also refers to the speed at which price changes follow a trend for a longer time. Expressly, the inertia of the price of securities may continue rising or falling for a period. Barberis et al. (1998)believe that the conservatism bias proposed by Edwards (1982) is a mental habit in making decisions that people tend to have difficulty acquiring new information due to persisting in their original idea or anticipation, leading individuals to underreact to the new information that makes the stock price show momentum effect. The most significant negligence of modern finance theories is cutting individual factors out of consideration. Additionally, Tomer (2007)concludes that behavioural economics and finance theories differ in that they are much more broad, flexible, tolerant, and connected than mainstream economic and finance theories. Human nature determines how individuals make decisions or take actions, which may have a huge impact on the market, and that humans will never make every decision rationally, as standard financial theory assumes.

Traditional financial theorists neglect the truth that most investors are humans, whether they are trained economists or unsophisticated investors. Behavioural economics and finance try to supplement mainstream theories to clarify the conundrums that trouble classical economists. The paragraphs above discussed the debate between traditional economics and finance. The next section describes the history and development of behavioural economics and finance.

## 2.4 The History of Behavioural Economics and Finance

The development of behavioural finance has faced significant doubts and rejections from the Neoclassical theory school. Despite these challenges, behavioural finance represents a new progressive suggestion in traditional theories.

It has been over 200 years since the concept of behavioural economics was first proposed. George Stigler recognizes that there is nothing entirely novel in the field of economics, as it has already been discussed by Smith. Thus, it is not surprising that Adam Smith is considered a pivotal figure who laid the foundation for behavioural economics in the history of economic research(Thaler, 2016).

Another human trait is practicality. It can be interpreted that men may be more attracted by the interest of today instead of the interest in the long future, no matter if the future interest exceeds present interest(Thaler, 2016).

The importance of behaviours is also observed by Keynes (1937); he explicates this circumstance in his famous magnum opus, *The General Theory of Employment, Interest, and Money.* He emphasizes that people's behaviour may be constrained by convention. A businessman often adheres to the practice of hiding his doubts and presenting himself as though he has a definite and accurate prediction. Typically, he must make decisions based on an average of various estimates, often erring on the side of caution for practical purposes. As a result, he is frequently compelled to make assumptions about long-term changes—about which he has limited knowledge—before he can confidently determine whether a particular strategy will be profitable.

Accordingly, research turns to the connection between the financial market and sociology and psychology, which includes the role of human activities and their behaviours or emotions in decision-making and investment progress. Shiller (2003)documents that the fields of behavioural finance developed because researchers have witnessed so many anomalies, exposing theoretical models to capture significant fluctuations. It is easy to understand why so many inexplicable anomalies exist in the economics and finance fields. Economics and finance are typically social sciences that research the routine operations of society. However, the establishment of society is based on human trading activities. Consequently, it is implausible and imprudent to disregard human factors. Pareto (1909)admits that one of the most significant factors is psychology, because it is the cornerstone of social science. However, he tried to eliminate psychology from economics research with his claim that people's preferences are foundational to economics but the cause of them is not important. Clark (1918)describes that economists attempt to ignore the role of psychology, but economics cannot ignore human nature because it is a subject with humans involved.

However, economists have often overlooked the factor of human behaviour, despite compelling evidence highlighting its significance. They tend to exaggerate the concept of the 'economic man' and downplay the influence of human nature. The field of mainstream economics has largely embraced their studies, exerting dominance in economics and finance for decades, starting from the early 20th century until the present. Although there have been numerous innovative scholars who have identified issues within mainstream economics, their voices have been overshadowed.

In addition, Fama (1998)claims that behavioural finance models can only explain the limited anomalies they were designed to explain; the efficiency market model has better performance on long-term return continuation than behavioural finance models. However, the behavioural finance model highlights the neglection of efficient market theories, and Fama does not answer why efficient market theories fail to explain the anomalies that are explained by the behavioural finance model. He also comments that behavioural economics pioneer Richard Thaler's research is interesting even though it lacks value because it does not provide a complete alternative model that offers better predictions.

From Fama's perspective, despite its insights, behavioural finance is primarily a critique of existing theories, not a constructive theory in its own right. Therefore, he regards behavioural economics as a limited subject, but he also neglects the shortage of mainstream economics, which is not as practical as behavioural economics. Nevertheless, Thaler's studies obtain explicit evidence that the flaws of human nature have a significant impact on economics as long as this area requires human involvement. The middle of the 20$^{th}$ century was an age of mainstream economics, which attracts scholars' attention. However, a conception regarded a cornerstone of behavioural finance known as 'bounded rationality' was discovered by Simon (1955); it suggests that most market participants' cognitive abilities are limited due to their limited knowledge of obtaining and dealing with information. This ability has an enormous effect on their decision-making process when they proceed with investments. Unfortunately, the concept was not valued at the time. This theory is the primary opposing opinion against the belief that investors are fully rational according to traditional theory. However, this phenomenon has been noticed before. In the 1970s, Daniel Kahneman and Amos Tversky tackle these problems by discussing human factors and redirecting people's attention from complex numbers or models to fundamental human behaviours. Tversky and Kahneman (1974)describe that there is a systematic bias in people's judgment-making processes. Their work is an important step for behavioural economics, using psychology as an entry point to research economic areas, and has enlightened future behavioural economist. Subsequently, behavioural economics and finance started to thrive, gaining famous researchers such as Richard Thaler and Robert Shiller, who each won the Nobel Prize; this was seen as a strong endorsement by the authority of the economic academic world.

Understanding the conflict and debate between school economics and finance and behavioural economics and finance, human consideration is an important difference among them. However, the human factor is invisible and difficult to capture; thus, the next section will discuss the method for monitoring the human factor in the market.

## 2.5 The Evaluation of Investment Behaviour

The previous section describes the significant effect of human factors or behaviours on the financial and economics area, derived from the field known as behavioural finance. Previous studies in this field use investor sentiment to represent the human factor and capture asset pricing or stock return anomalies(Brown et al., 2003; Brown and Cliff, 2004; Baker and Wurgler, 2006; Lemmon and Portniaguina, 2006; Baker and Wurgler, 2007).

However, before examining the evaluation of investing behaviours, it is crucial to investigate how these behaviours are formed. Imagine this story: the stock market experiences a sharp decline and decreasing trend for a period, with investors escaping the market as soon as they cannot afford it. It is easy to understand that a household investor may avoid unnecessary risk by depositing money when the interest rate is high rather than invest in a vague market.

Speculators regard the information spread by newsletters as an important factor in deciding whether they have made the right selection for their portfolio. Noise traders, who use noise information to determine when to invest, may take actions against rational economic principles(Kyle, 1985). In addition, the power of news and media plays a significant role. As Shiller (2016) documents in his book *Irrational Exuberance*, media and news easily influence people's activities, allowing them to omit the importance of precise quantitative analysis.

Additionally, numerous factors influence investment decisions, and dividends are among them. Investors often favour stocks that offer a higher likelihood of generating profits, as human nature tends to desire quick results and successes, especially in terms of financial gains. However, it is important to note that dividend allocation is subject to the discretion of managers. Shiller (1990)states that when the managers believe the biased optimistic mood fashioned in the system, they may increase dividends. Therefore, the stock value may increase as a result of investors finding that the dividend may pay them back quickly. At which point, they will rush to purchase that stock regardless the true root of the activity, which is fake optimism that may crush them in the future. Liquidity refers to the ability of an asset to be transferred into cash without impacting its original price, and Kyle (1985) defines it as the opposite of price sensitivity, a factor that cannot be overlooked during decision-making. The market will require new blood to start new investments, and liquidity may be a pump for that to happen. Liquidity helps new investments or projects survive their difficult initial period, and it looks like liquidity encourages people's willingness to take risks. This is explained by the model of De Long et al. (1990). It is not liquidity that drives people to take risks; it is the demand of noise traders that pushes the liquidity to a higher level. Sentiment, which is irrational expectation about the future market, drives the demand of noise traders to increase. The more noise traders misjudge, the higher the liquidity caused by their irrational demand. People tend to run a risk when investor sentiment stays high, and liquidity increases. Therefore, liquidity is a measurement of investor sentiment according to Baker and Wurgler (2006).

Stock prices also suffer from overreaction; stock prices have the tendency to react excessively to news or information and then later correct as the market realizes the initial reaction was too extreme. Kleidon (1986)believes that the correlation between stock price movement and earning changes shows the pattern of overreaction. In addition, De Bondt and Thaler (1995)identify the overreaction effect in the investment market, for example, professionals have an overreaction bias to news. This circumstance fits the classic patterns of investor sentiment confirmed by Baker and Wurgler (2006). The stocks increase rapidly, stay in bubble industries, and attract more attention from speculators, generating optimism. However, according to Black (1986), traders who utilize irrational noise as a signal for trading have the potential to earn higher returns compared to rational investors.

The phenomena above derive from behavioural aspects that require measurement, such as investor sentiment. Measuring the behavioural aspects could reflect cognitive biases, like loss aversion or overconfidence, and help understand market anomalies, like the momentum effect. Therefore, it is essential to explore methods for measuring investor sentiment that serve as widely accepted indicators since it is an abstract concept that exists in financial markets. What is sentiment, exactly?

Baker and Wurgler (2006)state that sentiment is difficult to monitor directly due to it depending on the patterns of investors' behaviours, including their psychology. A considerable volume of articles has studied investor sentiment and define it as different investors with different backgrounds holding various attitudes and anticipations toward their portfolios and future market movement(Baker and Wurgler, 2000; Singhvi, 2001; Brown and Cliff, 2004; Baker and Wurgler, 2006; 2007). However, it requires investigation since it has a significant impact on investors' choices and decisions. Nevertheless, there seems to be a debate about whether investor sentiment only influences individual investors or whether institutional investors are immune because of their access to information compared to individual investors(Chakravarty, 2001; Sias et al., 2006). In addition, Choi and Sias (2012)state that institutional investors may differ from individual investors because of their size and sophistication.

Although this argument remains controversial, mounting evidence suggests that institutional investors also suffer the impacts of investor sentiment. DeVault et al. (2019)state that institutional investors, because of their investment styles concerning risk management, reputation maintenance, or momentum trading strategies, tend to be sentiment traders by causing the sentiment-induced mispricing. First, they examine the relationship between sentiment metrics and individual investors' demand shocks and find that sentiment metrics capture the demand shocks of institutional investors rather than individual investors. This implies that institutional investors drive the sentiment-induced mispricing. Second, the authors find that differences in investment styles across types of institutional investors help explain institutional sentiment trading. Institutions with risk-averse investment styles, such as banks, insurance companies, pension funds, and unclassified institutions, contribute little to aggregate institutional sentiment trading as they tend to avoid holding and trading risky stocks. Institutions like mutual funds, independent advisors, and hedge funds are more willing to hold risky stocks and are more sensitive to lag performance, indicating that risk management and reputational concerns influence their sentiment trading. Additionally, they observe that transient institutions (those with high turnover and small stakes in individual companies) strongly engage in sentiment trading. This further supports the claim that institutional investors, due to their investment styles and strategies, are more likely to be sentiment traders and contribute to sentiment-induced mispricing.

In addition, according to Schmeling (2007), institutional sentiment might follow individual sentiment; institutional investors' sentiment may decrease when they believe individual sentiment is more optimistic, and institutional sentiment may increase if they expect individuals to become more optimistic because they realize that the price ought to be pushed higher by noise traders.

Hence, mutual funds are a good example of an institutional investor. There are previous studies related to mutual funds and investor sentiment(Ben-Rephael et al., 2012; Chiu and Kini, 2014; Hudson et al., 2020). It is essential to understand portfolio decisions for mutual funds before investigating sentiment's impacts. The portfolio decision of a mutual fund is generated in a few ways. The fund manager is usually responsible for placing orders and buying or selling specific stocks or bonds from the portfolio. Some smaller funds also have the lead manager handle marketing and backoffice activities. Larger mutual funds frequently have a staff of analysts, traders, and other personnel who watch the markets, conduct deals, and perform other activities for the fund. This support team is crucial to ensuring that the fund runs efficiently and profitably. However, it is the main mutual fund manager's responsibility to oversee the portfolio's general direction. Last, the manager decides what it will own and when. Mutual funds are sometimes operated by a committee. Principal fund managers gather opinions from each other, and stocks are chosen through a vote. A multi-manager fund is another typical approach. Each management team is assigned a proportion of the fund's assets to manage and is only accountable for those funds. A single lead manager will select who will be in charge and how much of the fund's assets will be assigned to them. Mutual funds that take investor sentiment into consideration when setting investment strategies may increase performance. For example, according to Shah and Baser (2022), investor sentiment can significantly influence the performance of mutual funds. This is particularly true around the turn of the month, when investor sentiment tends to be more positive, and fund managers take advantage of this phenomenon by adjusting their investment strategies accordingly. For example, they can increase their exposure to riskier assets at the end of the month, when investor sentiment is likely to be more positive, and reduce their exposure at the beginning of the month, when sentiment is likely to be more negative.

## 2.6 Different Measures of Investor Sentiment

There are mainly three methods to value sentiment: survey-based indicators, market-based indicators, and internet-based indicators. Sentiment indicators refer to the data that shows qualitative evaluation and prediction, including individuals' sentiment and subjectivity. For example, survey-based indicators such as the Investor

Intelligence Survey and the University of Michigan's surveyed consumer sentiment index, which are used by a large volume of scholars including Fisher and Statman (2000), Lemmon and Portniaguina (2006) and Schmeling (2009), are employed as a replacement sentiment measurement in the robustness check.

In addition, market-based indicators are a significant metric for research sentiment. Market-based indicators are divided into two kinds: single indicators and composited indicators. Single indicators refer to the closed-end fund discount rate or volatility index, and composited indicators refer to the Baker and Wurgler index, which includes NYSE measured trading volume, the closed-end fund discount, the number of IPOs and their first-day returns, the dividend premium, and the new issue equity share.

Additionally, recent attention has focused on the role that media such as news, social media platforms, or search volumes played in the market sentiment fluctuation (Tetlock, 2007; Bollen et al., 2011; Zhang et al., 2011;Bollen et al. (2021). Therefore, this section will introduce and elaborate upon the three types of sentiment measurements and compare their advantages and disadvantages. These measurements are presented in Figure 2–4. And the following sections will introduce each of them.

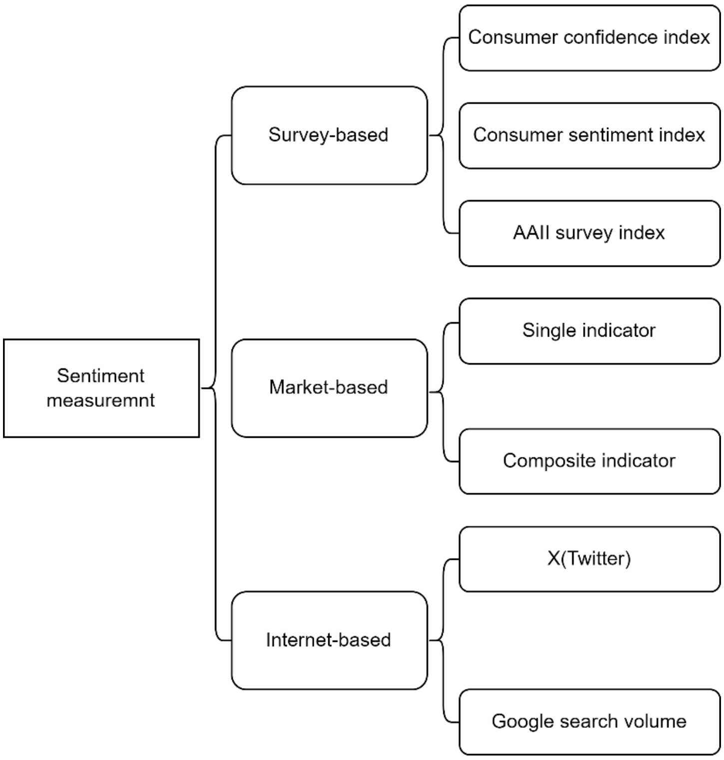



**Figure 2–4. Measures of investor sentiment**

## 2.6.1 Survey-based sentiment proxies

The first method is survey-based measures, which encompass several well-known indices. These indices can be classified into global range and regional range indices. Examples of regional indices include the consumer confidence index published by the Conference Board of the US, the consumer sentiment index (CCI) conducted by the University of Michigan, and the American Association of Individual Investors Survey (AAII). As for global indices, the Organization for Economic Co-operation and Development (OECD) Confidence Index is a notable example, covering 33 OECD member countries, six non-member economies, and eight zone aggregates.

Professor George Katona at the University of Michigan started the consumer sentiment index in the late 1940s and turned it from an annual investigation to a quarterly survey in 1952, then, in 1978, to a monthly survey. Katona's research is pioneering because it focuses on combining psychological research with economics and views economic changes through psychological perspectives long before the establishment of behavioural economics and finance. For example, Katona (1953)calls for psychologists to pay attention to the theory of rationality when researching individual behaviours concerning decision-making and problem-solving.

The consumer sentiment index has a smaller sample size than the consumer confidence index but a more detailed questionnaire, five questions on participants' thoughts about the business condition, economic situation, whether it is a good time to buy major household items, and their personal financial situation in the past and future. This index employs the data from 1966 as its benchmark value of 100. The investigation proceeds with monthly telephone interviews, and the number of samples is 500. There is an initial midmonth disclosure from approximately 250 samples collected early in the month. The final data from the full sample is available at the end of the month.

The formula is as follows:

$$CSI = \frac{R_1 + R_2 + R_3 + R_4 + R_5}{6.7558} + 2$$

To calculate the consumer sentiment index, a relative value is computed first through the numbers of positive replies less the number of negative replies and plus 100 for the five questions. Each value is rounded to the nearest whole number, and the sum of them is divided by the 1966 base period total of 6.7558, plus a constant to modify the sample design from the 1950s, which equals two. *CSI* refers to the consumer sentiment index, and *R* refers to the relative value of each question.

The consumer confidence index began in 1967, and it regards the data of 1985 as a benchmark equal to 100 because it is neither a peak nor low position. The survey, refreshed once a month, calculated the basic attitudes and expectations of 5000 individual household consumers toward the current and future movement of the economy. In addition, views about the current situation comprise 40% of the total, and the remaining 60% concern anticipation of the future. The equation is as follows:

$$R_t = \frac{P_t}{P_t + N_t}$$

$$I_t = \frac{R_t}{B_t}$$

$$CCI = \overline{\sum I_t}$$

where *R* equals relative value; *t* refers to the question's serial number; *P* refers to the number of positive answers; *N* refers to the number of negative answers; *I* refer to the indexed value; *B* refers to the benchmark value of 1985, and *CCI* refers to the value of the consumer confidence index.

The calculation of the consumer confidence index uses the total number of positive answers divided by the sum of negative and positive answer percentages and multiplying by 100, which gives the relative value of each question. The same benchmark value from 1985 is divided to obtain an indexed value. Finally, the five indexed values are averaged to calculate the consumer confidence index.

The AAII survey has been conducted among AAII's members every week with the same simple questions since 1987. Members are asked about the market's condition in the next six weeks with three categories to choose from: bullish, bearish, or neutral. Approximately 100 survey questionnaires are distributed to AAII's members weekly; collected responses are analysed to determine the attitude of investors. If the results show too many investors selecting bearish, it very likely reflects the market lows. If there are many bulls, it indicates that the market sentiment will reach a high point soon.

Though some researchers, such as Fisher and Statman (2000), use it as a useful sentiment index, the AAII survey has a limitation.; the number and size of its questionnaires is quite small, because they only ask their members. Fisher and Statman (2000)only use the survey results to represent the sentiment of small investors, not the whole scale of the group. In addition, Qiu and Welch (2004)state that the AAII survey has a low ability to follow the changes of the same individual at different times, and it gives its participants a high degree of freedom to answer or not answer the questions. Hence, self-selection bias appears, in which undesirable or abnormal situations caused by the characteristics of the people cause them to choose themselves in the group and generate a biased sample with nonprobability sampling. Therefore, the AAII's survey is not a perfect index to measure sentiment, and this thesis will not include it. By contrast, the consumer confidence index and consumer sentiment index are proven to properly reflect investor sentiment(Fisher and Statman, 2003; Bergman, 2008; Schmeling, 2009).

Overall, survey-based indicators provide direct and explicit measures of investor sentiment by asking individuals about their views on the economy and markets. This method offers valuable insights into the psychological aspects of market behaviour, particularly in understanding how sentiment influences investment decisions. However, survey-based measures are not without limitations. The sample size limitations of surveys may not represent the entire investor population Additionally, self-selection bias can occur when respondents choose whether to participate, potentially skewing results. The timing and frequency of surveys may also limit their ability to capture rapid changes in sentiment. Despite these limitations, survey-based indicators like the Consumer Confidence Index and Consumer have been shown to effectively reflect investor sentiment and are thus included in the empirical analysis(Bergman, 2008; Schmeling, 2009).

### 2.6.2 Market-based sentiment proxies

In addition to the direct sentiment measurement (survey-based), the indirect sentiment indicators (market-based) deserve to be discussed. The market measurements can be divided into single indicators and composite indicators. As for single indicators, many have been researched to observe their relationship with sentiment, including the market liquidity aspect measures such as illiquidity, bid-ask spreads, the price impact of trades, and the trading volume as measured by NYSE turnover, as well as other aspects like mutual fund redemptions, the dividend premium, the closed-end fund discount, and the number and first-day returns on IPOs. The relationship between market-based measurements can be observed in Figure 2–5.

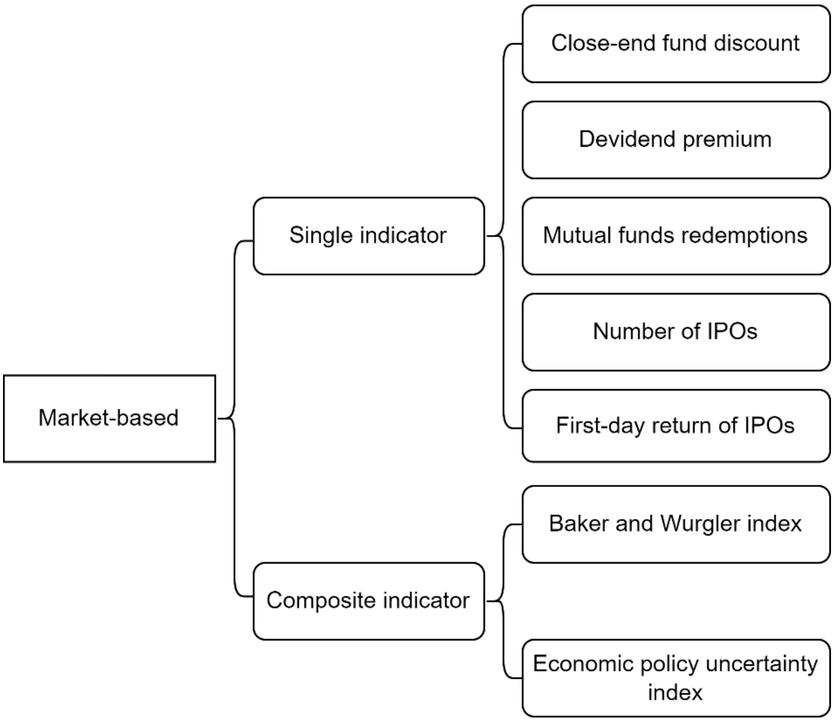



**Figure 2–5. Market-based measurements**

#### 2.6.2.1 Single investor sentiment proxies

Market liquidity refers to the ability to buy or sell an asset without causing a significant change in its price. It is a market characteristic that describes the degree to which an asset or security can be quickly bought or sold in the market at a price reflecting its intrinsic value. An increasing number of empirical studies indicate that the performance of stocks can be predicted by market liquidity (Amihud and Mendelson, 1986; De Long et al., 1990; Brennan and Subrahmanyam, 1996; Brennan et al., 2012). Baker and Stein (2004) suggest that some investors regard this as a signal and believe they are more well-informed than others, which makes them overconfident or remain at a high sentiment level to overweight an asset, without knowing a lower expected return is awaiting.

Sentiment refers to investor's combination of belief and desire toward the future market, which could be bullish or bearish. When sentiment is high, irrational people tend to rush into the market by selling their assets to cash out and reinvest in another investment, which leads to increased liquidity. The result of investor sentiment boosts market liquidity, as confirmed by Liu (2015). The increasing sentiment informs investors that the sentiment of the market is increasing, leading them to believe the market return will rise in the next period. Investors regard indicators like market liquidity, which has the power to predict future returns, to make their expectations, and even higher liquidity is caused by the irrational demand of noise traders. Hence, market liquidity is an important measure of sentiment. Brown and Cliff (2004)and Baker and Stein (2004) detect the connection between market liquidity and sentiment. Hence, the measure of market liquidity can also measure sentiment.

Market liquidity can also measure sentiment, including illiquidity, price impact, bid-ask spreads, trading volume, and turnover ratio. Therefore, illiquidity illustrates the reaction of daily price changes linked with a one dollar change in trading volume. Thus, Amihud (2002)considers illiquidity a market liquidity measure; it also measures price impact.

The price Impact of trades indicates there is a link between the incoming trade orders and the price fluctuation after the trades. For example, the price will be pushed up by a buy trade, which is easy to observe, and the investors' second purchase will be more expensive than their first due to the impact of their first transaction (and vice versa for sell trades). In addition, according to the research by Glosten (1989), Kyle (1985), and Easley and O'Hara (1987), the liquidity effects can be detected by the price impact of trades. However, Amihud (2002)states that illiquidity measures the price impact of trades and may lead to an autocorrelation problem between them, which refers to the degree of correlation of the same variables across different observations in a dataset. Hence, price impact is not included as a market sentiment indicator in this thesis.

Furthermore, the bid-ask spread is a sentiment measurement through market liquidity. It is the amount that the bid price is lower than the ask price, and it indicates the difference between the lowest price that a trader is willing to sell and the highest price that a buyer is willing to purchase. When the spread is equal to zero, then it is a frictionless asset, which suggests that a lower bid-ask spread indicates higher liquidity of the asset. Many previous studies document that the bid-ask spread has the ability to demonstrate market liquidity, including work by Amihud and Mendelson (1986).

However, this measure can only be used annually, and it is not available in many stock markets. Even if it were, it does not cover a long enough period. In addition, the bid-ask spread is regarded as a noisy measure of illiquidity by Lee et al. (1993) due to the fact that it cannot take the necessary trades into account, such as many small trades occurring in the spread while many large trades occur outside of the spread. Hence, the bid-ask spread (effective or quoted) is not employed in this thesis.

There is a better method to value liquidity, and that is trading volume. Trading volume is a prescribed security's total number of exchanged shares. This could be scaled to any type of security exchanged during any trading day. Commodities such as bonds, options, stocks, and futures contracts can be measured by trading volume. Brennan and Subrahmanyam (1996)describe trading volume as a crucial indicator of market liquidity. When the trading volume increases, more people trade on specific assets, which suggests the willingness to trade increases. Baker and Stein (2004)emphasize that irrational investors are more willing to trade, which boosts liquidity, when they feel optimistic and gamble on the rising stocks, rather than when they feel pessimistic and gamble on the decreasing stocks. Therefore, trading volume can show the movement of liquidity, and Baker and Wurgler (2007) view it as a liquidity aspect of sentiment measurement.

The turnover ratio is another measurement of investor sentiment; it is the replacement percentage of portfolios or funds during a certain period. Additionally, it is a measure of trading volume because a high turnover rate suggests investors tend to exchange their portfolios, for example, from holding mutual funds to holding stocks. This results in a higher turnover ratio and indicates a higher willingness to trade as well as a higher trading volume. Therefore, many researchers see turnover ratio as an important sentiment metric. Scheinkman and Xiong (2003)demonstrate that turnover rate is a significant proxy for investor sentiment when the market is overconfident.

Baker and Wurgler (2006)employ it as a measurement of sentiment.

Except for market liquidity measurements, other single indicators can monitor investor sentiment. Mutual fund redemptions are one of them, which refers to investors selling or redeeming the funds to withdraw their capital as a kind of mutual fund flow. Malkiel (1977)regards mutual funds as a ‘general sentiment measurement’. The reason for an increase in mutual fund redemptions could be an abrupt rise in the stock market. At this point, investors see an opportunity for profit and enter the stock market heavily. Why can mutual fund flows predict investor sentiment? This is because the holders of mutual funds are mostly unsophisticated and less informed investors who have given their trust to their mutual fund managers. Such investors are easily impacted by the price movement of the market. Therefore, Warther (1995)states that mutual fund flow is a proper place to look for unsophisticated investor sentiment indicators.

Another method to identify investor sentiment in the market is the closed-end fund discount rate. Closed-end funds are types of mutual funds that possess other publicly traded securities. The difference between them and open-end funds is that the number of shares that closed-end funds issue and trade on the stock market are limited. Investors can only sell their shares to other investors and cannot redeem their shares with the fund, just as with open-end funds. However, closed-end fund shares are usually traded at a price unequal to the market value of the assets the fund holds. It might sell at a discount, though it could sell at a premium to their net asset values; this is a conundrum known as the ‘closed-end fund puzzle’, discovered by Zweig (1973) andDe Long et al. (1990). Weiss (1989) also observes a mysterious condition: investors tend to be more willing to purchase younger funds and ignore the fact that younger funds usually underperform. Berk and Stanton (2007)wonder the result if investors choose to purchase a closed-end fund at its IPO price, which is possible to fall to a discount, while investors understand they can buy an open-end fund that guarantees trading at market price in the future. Investor irrationality could be the reason behind this situation, which has a link with market sentiment, and the discount rate will increase if investor sentiment becomes pessimistic. Lee et al. (1991)propose that the investor sentiment approach can explain the closed-end fund puzzle by showing that the changing of investor sentiment makes the risk of the fund higher than purchasing those securities directly. The closed-end funds will sell at a discount to induce new investors to buy them; Lee et al. (1991)find that using the closed-end discount rate can represent investor sentiment on the reverse. However, their finding is challenged. Qiu and Welch (2004)state that the result of the closed-end discount rate has no correlation with sentiment measurement and only works in demonstrating a small stock excess return. They also state that the closed-end discount rate is not a proper proxy for investor sentiment due to its timeliness; it has not worked since 1985. Due to this, this sentiment proxy will be excluded.

Baker and Wurgler (2004)Baker and Wurgler (2004)The next measure to discuss is the volume of IPOs and their first-day returns. Ritter (1984; 1991)observes that there was a hot period of IPOs during which their volume increased rapidly. According to Baker and Wurgler (2000) and Ljungqvist et al. (2006), companies went public during this overoptimistic period to purposely be overpriced. A high volume of IPOs indicates that more companies are willing to go public, which suggests sentiment is high(Lee et al., 1991). Overenthusiasm firms have more incentives to go public when investors are exuberant, with the aim of taking advantage of those noise investors to extract a surplus from them. Firms will choose this period because it has a greater possibility of achieving a higher price. In addition, the enthusiasm of investors will fade, and the number of sentiment investors is limited; therefore, firms are likely to compete in the IPO market and regard those investors as a kind of resource. Hence, if there is a hot period in the IPO market, the sentiment during this period is likely to be high as well.

There is evidence that IPOs tend to be underpriced and achieve a high first-day return, but then a long-run underperformance follows, making IPOs seem overpriced. Ofek and Richardson (2003)document large declines in the price of IPOs caused by burst bubbles after retail investors see high initial returns on the IPO shares purchased from institutions on the first day. Ritter (1991)finds evidence of unusually low performance in the long term after high first-day returns. Purnanandam and Swaminathan (2004)present evidence that issues have higher first-day returns and lower returns in the long run by virtue of their overpricing compared to a fair value.

This phenomenon may result from the optimistic expectations or desires of investors or represent sentiment and cause offerings to be priced above their fundamental value. After the expectations of investors turn realistic, the price meets a reversion until the price reaches the fundamental value. Cornelli et al. (2006)suggest that investor sentiment may be the reason behind the high first-day returns due to sentiment investors, whom they call ‘grey market investors’, who are willing to afford a price above the fundamental price when they are overoptimistic, which pushes the aftermarket price up. Hence, first day returns on IPOs and the volume of IPOs could be proxy measures of investor sentiment.

In summary, Market-based indicators are often favoured for their ability to reflect actual investor behaviour in real-time, offering a more dynamic and continuous measure of sentiment compared to surveys. For instance, when sentiment is high, firms are more likely to undertake IPOs, anticipating strong investor demand and higher valuations(Ljungqvist et al., 2006). The number of IPOs is thus a direct reflection of market participants' willingness to invest in new and often speculative ventures, making it a pertinent measure of sentiment in equity markets. Nevertheless, the primary limitation of market-based indicators lies in their indirect nature; they infer sentiment from market behaviours rather than measuring it directly. This can lead to misinterpretations, as these indicators are also influenced by factors other than sentiment, such as macroeconomic conditions or regulatory changes(Da et al., 2015).

#### 2.6.2.2 Composite sentiment proxy

Several single indices for investor sentiment have been introduced. Is it more efficient to use them together? Will one take care of the other's shortcoming? Various scholars have combined different sentiment indices in their research on investor sentiment measurement. For example, the BW Index is formed by Baker and Wurgler (2006; 2007). This index consists of six different market-based proxies that can show the movement of investor sentiment in the market: the NYSE turnover rate, dividend premium, closed-end fund discount, volume of IPOs, the percentage of stocks in new issues, and the first day return of IPOs. The proxies form a composite index with the first principal component analysis. Baker and Wurgler (2006)believe that it is hard to find a perfect indicator without any controversy to measure investor sentiment; thus, a composite index is useful. For instance, Qiu and Welch (2004)describe that closed-end fund discounts might not represent sentiment precisely due to the fact that they can be impacted by other factors such as transaction costs or agency costs. Da et al. (2015)argue that the market-based indicators are not pure sentiment indicators. The shortcomings of other economic factors cannot be ignored; although, market-based indices may be more readily accessible at a higher frequency than other indices.

Therefore, Baker and Wurgler (2007)prefer a purified sentiment index and conduct a composite index by combining the market-based measures to reduce the influence of other economic forces. Baker and Wurgler (2007)try to eliminate the influence of investment opportunities implied on the number of IPOs; they regress each proxy on a group of economic indicators like industrial production growth, employment growth, real growth in durable and nondurable goods, and the recession index. Next, they use the residuals from those regressions to form their sentiment measure. Hence, the BW Index has become an influential and widely known sentiment indicator. The formula (Baker and Wurgler, 2007) is as follows:

$$SENT = -0.23CEFD+0.23TURN+0.24NIPO+0.29RIPO -0.32PDND+0.23S$$

where *CEFD* refers to closed-end fund discount; *TURN* refers to turnover ratio; *NIPO* refers to the number of IPOs; *RIPO* refers to the first day return of IPOs; *S* refers to the equity share, and *PDND* refers to dividend premium.

Composite indicators, such as the Baker and Wurgler (BW) Index, aggregate multiple single indicators to create a more comprehensive measure of sentiment. The BW Index, which includes variables like NYSE trading volume and IPO activity, attempts to filter out the noise by combining various market-based proxies, thus providing a more stable and reliable sentiment measure(Baker and Wurgler, 2007). Many scholars accept the BW index as an efficient method to value sentiment in their research, such as Yu and Yuan (2011), Mian and Sankaraguruswamy (2012) and Huang et al. (2015). However, the BW Index has flaws, like the fact that it is difficult to measure sentiment on a worldwide scale. As Chang et al. (2012) state, it is difficult to find a counterpart for many countries, especially those developing countries. Either the data is not available, is only recorded for a short period, or is prohibited and very difficult to construct. But the BW Index is employed due to its broad acceptance in the literature and its ability to integrate multiple dimensions of market sentiment, making it a valuable tool for analysing investor behaviour in various market conditions.

### 2.6.3 Internet-based sentiment proxies

The final sentiment measure that has gained attention recently is observing the investor's sentiment level using internet data from a search engine or social media platform. Internet-based sentiment indices are constructed using data sourced from online platforms such as social media (e.g., Twitter) and search engines (e.g., Google). These platforms provide vast amounts of real-time data, which can be analysed to gauge public sentiment on various topics, including financial markets. For example, Twitter's API allows researchers to collect tweets based on specific keywords. This data includes the text of the tweets, timestamps, user information, and metadata such as the number of retweets and likes. Researchers might focus on collecting tweets related to specific stocks, companies, or economic terms(Bollen et al., 2011; Zhang et al., 2011). Moreover, Google Trends is a popular tool for collecting search volume data for specific terms over time. Researchers can use this data to track changes in search interest, which can serve as a proxy for public interest or concern about a particular topic(Choi and Varian, 2012).

Prior research focuses on the relationship between internet messages and stock market situations, like Antweiler and Frank (2004) and Gilbert and Karahalios (2010). For example, formerly and still colloquially known as Twitter, X (hereafter Twitter), is a popular social media platform with many users who share their thoughts, follow other users they like, retweet other users' posts, and communicate with each other. Expressly, Twitter is a reflection or prediction of users' activities or moods.

Accordingly, Zhang et al. (2011)collect the data of tweets and retweets from March 2009 to September 2009 to examine the correlation between the number of emotional Twitter feeds, such as those describing hope, fear, or worry and the stock market. As a result, they find that there is a negative correlation between the number of emotional feeds and the Dow Jones price. However, they test the relationship between market and Twitter. Mao et al. (2011)discover that Twitter sentiment has a significant effect on market returns.

Furthermore, search engine data is the next internet-based sentiment indicator. The searching activities of individuals may show their focus or beliefs. Imagine this story: an investor hears a whisper about which company may make a profit or suffer a loss in the near future, and he puts money into that firm by coincidence. It is easy to understand that he will worry whether his investment is wise. However, he has limited methods to obtain useful information, and he might use the internet to inquire whether the news he has been told is true or false. Choi and Varian (2012)were among the pioneers in using Google search data to predict economic indicators, showing that search trends could effectively forecast consumer behaviour and market activity. Choi et al. (2002)state that investors use the internet as a channel of news and transactions, which may generate a large ‘web effect’, such as doubled trading frequency or an increased portfolio turnover ratio. Therefore, the data from the search engine might illustrate the attitude or sentiment of individuals toward the market or a single security. For example, Das and Chen (2007)extract an index from the internet to measure small investor sentiment.

Beer et al. (2012)propose a proper indicator that could capture a better view of investor sentiment by researching French investors, extracting the search volume from Google Trends, and examining its influence on the stock market. In addition, Gao et al. (2020)construct a global sentiment using Google search behaviours that covers 38 countries during the market period 2004–2014 and prove it has a negative correlation with market returns. Therefore, search engine data is a reliable measure of sentiment.

Internet-based sentiment indices are bounded to be several limitations. One key issue is that of data representativeness. Social media platforms and search engines may not provide a representative sample of the overall investor population. Users of these platforms tend to skew younger and may not reflect the views or behaviours of older or more traditional investors. Additionally, sentiment expressed online may be biased due to individuals are exposed to and reinforce similar opinions, leading to skewed sentiment measures. Twitter's user base, for instance, is not necessarily representative of the general population or the broader investor community. This can result in sentiment indices that are more reflective of the sentiments of a specific demographic rather than the entire market(Bollen et al., 2011). Furthermore, internet-based sentiment indices often carry a high level of noise and overshadow meaningful signals. Social media and search engines capture all sorts of user-generated content, much of it having no relevance to financial markets or investor sentiment. For example, trending topics on Twitter or spikes in search volumes may have nothing to do with financial events, so the data could be misleadingly interpreted to reflect investor sentiment. Zhang et al. (2011)highlight that while Twitter sentiment can predict stock market movements, the presence of unrelated or irrelevant tweets can dilute the quality of sentiment analysis.

### 2.6.4 Recent developments of investor sentiment analysis

The measurement of investor sentiment has evolved significantly with the advent of new technologies and methodologies. Traditional measures such as survey-based and market-based indicators have been complemented by innovative approaches, including Natural Language Processing (NLP), sentiment analysis using machine learning models.

Natural Language Processing (NLP) is a branch of artificial intelligence that enables computers to understand, interpret, and generate human language. In the context of financial markets, NLP has been increasingly applied to analyse large volumes of textual data from sources such as financial news, social media, and corporate reports. The objective is to extract sentiment or emotions that can be linked to market behaviours. Tetlock (2007)applied NLP techniques to extract sentiment from the Wall Street Journal's "Abreast of the Market" column, demonstrating that high levels of negative sentiment predicted downward pressure on market prices.

Machine learning models have been widely adopted for sentiment analysis due to their ability to handle large datasets and improve accuracy over time through learning from data. These models can be trained on labelled datasets to classify sentiment in financial texts or predict market trends based on sentiment scores. Das and Chen (2007)employed machine learning techniques to classify investor sentiment in online stock message boards. Their model was able to predict stock market movements by categorizing posts as bullish, bearish, or neutral.

While machine learning (ML) and natural language processing (NLP) techniques have brought significant advancements in analysing investor sentiment, these methods also face several limitations. These limitations can impact the accuracy, reliability, and interpretability of sentiment analysis in financial markets. Below are some of the key limitations. ML and NLP models rely heavily on the quality of the data they are trained on. In the context of investor sentiment, data from sources like social media, news articles, or financial reports can be noisy and contain irrelevant or misleading information. This noise can lead to inaccurate sentiment analysis. Bollen et al. (2011)highlighted that while Twitter data can be useful for predicting market movements, the platform also contains a significant amount of noise—irrelevant tweets or posts that do not contribute meaningfully to sentiment analysis.

Another significant challenge with ML and NLP models, especially deep learning models like neural networks, is the lack of interpretability. These models often operate as "black boxes," making it difficult for analysts to understand how the model arrived at a particular sentiment classification or prediction. This lack of transparency can be problematic in financial markets, where understanding the rationale behind sentiment analysis is crucial for making informed decisions. Loughran and McDonald (2016)emphasize that while advanced NLP techniques can be powerful, their opacity can be a drawback in finance, where stakeholders need to understand the underlying reasoning for predictions. This can limit the practical applicability of such models in high-stakes environments like financial trading.

Furthermore, investor sentiment is not static; it changes over time in response to new information and evolving market conditions. However, ML and NLP models can struggle with capturing these temporal dynamics, particularly if the models are trained on data from a specific time that may not reflect future conditions.  discusses the challenges of using sentiment models that do not account for temporal shifts in sentiment. Without dynamic updating, these models may become outdated quickly, leading to poor performance in changing market environments. Khadjeh Nassirtoussi et al. (2014)discuss the necessity of dynamic models in text mining for market prediction. The authors emphasize that static sentiment models, which do not account for the temporal evolution of sentiment, risk becoming outdated and less effective in changing market environments.

Overall, machine learning and NLP offer powerful tools for analysing investor sentiment, but they are not without limitations. Issues such as data quality, model overfitting, lack of interpretability, language context, and temporal dynamics can all impact the effectiveness of sentiment analysis. Therefore, this thesis will focus on traditional investor sentiment indices rather than using the techniques mentioned above.

## 2.7 Incorporating Prospect Theory in the Nexus Between Investor Sentiment and Portfolio Composition of Mutual Funds

I now turn my attention to prospect theory in the nexus between investor sentiment and mutual funds' portfolio composition. Prospect theory, developed by Kahneman and Tversky (1979), revolutionized the understanding of decision-making under risk by introducing concepts such as loss aversion and the value function, which are central to explaining investor behaviour. Unlike traditional expected utility theory, which assumes rational decision-making based on final outcomes, Prospect Theory emphasizes the psychological biases that influence individuals' decisions, particularly how they perceive gains and losses relative to a reference point.

The two main components of Prospect Theory are the value function and the probability weighting function. The value function is concave for gains and convex for losses, indicating that people are risk-averse for gains and risk-seeking for losses. This S-shaped function reflects the diminishing sensitivity to both gains and losses as their magnitude increases, with losses perceived more strongly than equivalent gains—a phenomenon known as loss aversion. Kahneman and Tversky (1979)showed that losses are typically twice as impactful as gains, a principle that has profound implications for financial decision-making, particularly in the context of investment strategies.

Loss aversion can influence mutual fund portfolio composition by affecting fund manager decision-making. Loss-averse investors may exhibit a preference for lower-risk mutual funds that are less likely to experience significant losses. Fund managers with a greater sensitivity to losses select portfolios with lower downside risk, increase their risk-taking in reaction to poor previous performance, and exhibit a rather stronger disposition effect. In addition, empirical evidence demonstrates that managers with higher loss aversion tend to generate less profit and are more likely to have their contracts terminated, underscoring the economic consequences of this behavioural bias.(Bodnaruk and Simonov, 2016).

The value function can influence mutual fund portfolio composition through its impact on investor risk preferences. According to prospect theory, individuals exhibit risk-seeking behaviour for losses and risk-averse behaviour for gains. This asymmetric risk attitude can result in suboptimal portfolio decisions, as investors may hold onto losing investments for too long in the hope of a rebound while selling winning investments too early to lock in gains(Shefrin and Statman, 1985). According to Giannikos et al. (2023), the behaviours of fund managers are more consistent with prospect theory than utility theory. They identify the deposition effect by proving that the trading cost decisions of managers exhibit ‘play it safe’ behaviour. Moreover, Yu et al. (2022)state mutual fund managers suffer from behavioural biases, such as dynamic loss aversion, which refers to the phenomenon in which an individual’s sensitivity to losses varies depending on the context, particularly in relation to past gains or losses, and it may influence their investment decisions. Furthermore, Gu and

Yoo (2021)propose the PT value, which is a measure of the attractiveness of an investment under prospect theory. Expressly, it is a measure of the degree of convexity or concavity in investors' preferences. A higher PT value indicates that the fund's investors are more likely to be loss-averse and may be willing to take more risks to avoid losses than to achieve gains, and the return distribution of mutual funds is more aligned with investors' risk preferences. Gu and Yoo (2021)demonstrate that mutual funds with higher PT values tend to have better future performance, which suggests that investors may be rewarded for taking on loss aversion. According to Gu and Yoo (2021), mutual fund managers may communicate the PT values of their portfolios to investors to differentiate themselves from competitors and attract more assets. Yu et al. (2022)state that mutual funds with a high PT value may attract more investment from naive and irrational individual investors, resulting in higher net fund inflows and possibly higher returns. All the finds above suggest that prospect theory is important to the portfolio composition of mutual funds.

After discussing the connection between prospect theory and mutual funds portfolio composition. The empirical link between investor sentiment and mutual fund portfolio composition is also well-documented. Ben-Rephael et al. (2012)examine mutual fund flows as a measure of investor sentiment, showing that funds experience higher inflows during periods of positive sentiment. This behaviour suggests that mutual fund managers are influenced by the prevailing sentiment, adjusting their portfolios to align with investor expectations. This alignment can lead to pro-cyclical behaviour, where managers increase risk exposure during bull markets and reduce it during bear markets, potentially exacerbating market cycles. Furthermore, studies such as Chue and Mian (2022) explore how sentiment affects mutual fund managers' stock-picking behaviour. Their research indicates that fund managers tend to reduce active stock selection during periods of high sentiment, staying closer to their benchmarks. This behaviour can be regarded as managers are not sensitive to positive sentiment shocks and challenges the conventional view that only retail investors are sentiment-prone, suggesting that institutional investors, including mutual fund managers, are also influenced by market sentiment. The reduction in active share during high sentiment periods can lead to increased mispricing and a higher prevalence of cross-sectional anomalies, as managers focus less on individual stock selection and more on aligning with broader market trends.

Therefore, incorporating a comprehensive understanding of prospect theory enhances the theoretical framework, providing a robust foundation for analysing the impact of investor sentiment on mutual fund portfolio composition. It is helpful to explain mutual fund portfolio composition or mutual fund managers' investment decisions and to identify better performing mutual funds and increase potential profits.

# 2.8 Conclusion

This chapter describes the development and history of behavioural economics and finance and explains the theories' popularity, despite the obvious failures and weaknesses of traditional theories. This chapter confirms the root of some behavioural finance views within economic development. The behavioural factors that seem invisible to traditional numeric calculations of traditional economics and finance are covered by one index: investor sentiment. Therefore, this chapter elaborates on evaluating this abstract index using different methods. The methods are classified by different data access methodologies, such as survey-based, market-based, and internet-based. Additionally, this chapter discusses the advantages and disadvantages of different measures. Survey-based measures are direct measures of investor sentiment, like the consumer confidence index and consumer sentiment index, which are easy to access, but the limitation is that some participants may not answer honestly due to personal reasons. There are also market-based proxies; some are single indicators, and some are composited. Nevertheless, the single proxy needs to be selected due to the number of indicators; this thesis chooses the number of IPOs, the first day return of IPOs, and the volatility index, because they have fewer criticisms than others. The closed-end discounted rate and bid-ask spread are excluded because of their respective flaws. The widely known composited proxy constructed by Baker and Wurgler (2006), the BW Index, is employed in this thesis as a measure of sentiment to determine whether the composite index carries more information than single proxies. Finally, the internet-based measurements, which are Twitter sentiment and Google search engine data, are employed as proxy measures to monitor individual investor sentiment.

Among the sentiment metrics mentioned, I choose the consumer confidence index, Google search volume, number and first-day return of IPOs, BW Index, and volatility index to continue my empirical research.

# Chapter 3. The Relationship Between Multiple-Based Investor Sentiment Proxies and Mutual Fund Portfolio Composition

## 3.1 Introduction

With a growing number of market anomalies arising that traditional economics cannot explain, the role of behavioural economics and one of its crucial indicators, investor sentiment, has attracted increasing academic attention. Many pioneering studies show that investor sentiment impacts investor behaviour, such as herding activities, including among institutional investors such as fund managers(Liao et al., 2011; Massa and Yadav, 2015; Hudson et al., 2020). However, this raises a question: how should institutional investors engage in investment activities? They may choose to change the composition of the portfolio they are holding or have constructed, for example, by selling or purchasing more specific securities. Hence, this raises more questions: will the investor sentiment level impact the composition of the portfolio? In addition, whether the change resulted from the price fluctuation or the increasing investment amounts of the underlying assets. This research using stock market index to as control variable o address this question in Appendix. In addition, which sentiment index has more information concerning mutual funds portfolio allocation? Previous research has not addressed this issue.

Nearly all research on investor sentiment concerns sentiment's impact on stock returns, fund performance, or considers mutual fund flows as an indicator of investor sentiment. Ben-Rephael et al. (2012) document that mutual fund flow can be used to estimate stock market sentiment in the US market. According to Indro (2010), net aggregate equity fund flows are related to the sentiment of both individual investors and newsletter writers. Frugier (2016) selects 46 stocks from the European stock market to create various portfolios to investigate whether returns differ when considering investor sentiment, instead of utilizing a mutual fund portfolio active in the market. Consequently, his research solely concludes how sentiment might impact a portfolio at a specific time point. Bu and Forrest (2021) and Bu (2021) only compare a limited number of sentiment measures reflecting mutual fund performance rather than portfolio composition. For example, the direct sentiment measure AAII may be more effective than the indirect sentiment measure BW Index in predicting stock and mutual fund performance. Chue and Mian (2022) present only two sentiment indicators in their study. They explore the impact of these indicators on mutual fund portfolio activeness, using ‘active share’—the percentage of a fund’s portfolio that deviates from its benchmark index, as the dependent variable. This metric reflects the efforts of fund managers to outperform the market. However, their study does not examine various benchmark indices or determine which indices most accurately reflect portfolio composition. Additionally, the relevance of their research is constrained by the sample period, which spans from the first quarter of 1985 to the third quarter of 2009.

In summary, this study contributes to the literature by focusing on the relationship between investor sentiment and portfolio composition. Moreover, few studies have concentrated on the connection between investor sentiment and existing portfolios in the market, like mutual fund portfolios. In addition, this chapter employs five multiple-based investor sentiment indices: a survey-based index, the consumer confidence index; an internet-based index, including the Google Trends search volume of key words representing negative sentiment, and the single market-based index representing the number and first-day return of new issued IPOs; the CBOE volatility index; and the Baker and Wurgler Index, a composite market-based index. Hence, these indices reveal whether investor sentiment impacts the professional portfolio composition of different mutual funds and determine which has the strongest relationship.

This chapter proceeds in four sections. The next section outlines the data sample and methodology. Section 3.3 discusses the empirical results. Section 3.4 demonstrate the robustness check. And the last section offers conclusion.

# 3.2 Data and Methods

## 3.2.1 Data

It is difficult to obtain sufficient and powerful data on the exact individual composition of portfolios, some of which is private personal information that is difficult to disclose. The selection of data periods and variables is based on their availability. Therefore, this study uses institutional portfolio data instead, which contains two dependent variables, to represent the portfolio composition. Portfolio composition is the percentages of capital held in different categories of assets by a mutual fund portfolio, like common stocks and cash. The dataset consists of 95 selected mutual funds in the US market from the Center for Research in Security

Prices (CRSP) of the Wharton Research Data Service database.[1] I exclude the mutual funds that are not consistent with the research sample period; for example, some funds are not established at the start date of the research period or shut down during the period. Excluding the duplicated data from the CRSP's mutual fund database, 95 mutual funds remain.

This study compares the effectiveness of various sentiment measurement methods in reflecting portfolio composition. This study follows the prior literature on investor sentiment for considering control fundamental variables(Smales, 2017), using the interest rate and unemployment rate to indicate internal market stability. For instance, the unemployment rate could be a crucial proxy for market health, and the interest rate may reflect the money policy. Additionally, the geopolitical risk index can be used to represent external market stability. It can reveal whether the country is facing systemic risks from a geographic or political aspect, such as warfare. In addition, He (2023) documents that geopolitical risk index may impact investor sentiment. Gu et al. (2023) state that portfolio need to construct hedging strategies against geopolitical risk. Because the geopolitical risk index may influence the return of portfolio(Zaremba et al., 2022). Therefore, it is necessary to include geopolitical risk index as control variable.

The data sample period ranges from 2007 Q3 to 2021 Q2, which includes the 2008 financial crisis and the COVID-19 recession. Meanwhile, the independent variable data are obtained from multiple sources; the number of IPOs and the first day

[1] The Wharton Research Data Service database website is https://wrds-www.wharton.upenn.edu/

return of IPOs are obtained from the dataset on Jay Ritter's website.[2] Other variables, such as the consumer confidence index, the CBOE volatility index, and the unemployment rate, are obtained from DataStream; the consumer sentiment index is obtained from the Surveys of Consumers of University of Michigan.[3] The BW Index is accessed through the dataset listed on Jeffrey Wurgler's website.[4] The short-term interest deposit rate used to represent the interest rate is obtained from the Organization for Economic Co-operation and Development (OCED) database.[5] The geopolitical risk index is obtained from the Economic Policy Uncertainty website.[6] Finally, Google Trends provides an internet-based sentiment indicator in the form of the Google search volume for the terms 'financial crisis' and 'economic recession'. The reason for selecting these indicators is presented in Chapter 2.

First, this study follows the standard procedures for variables in a time series analysis to conduct descriptive statistics, as shown in Table 3–1. The reason why stock investment and cash holdings can have 269% and -87.78% could be the strategies of mutual funds like leverage and short selling. Leverage allows a fund to borrow capital in addition to its equity. This borrowed capital is then used to purchase additional securities, like stocks, and make the value exceeding 100%. While short selling entails selling borrowed securities with the obligation to repurchase them in the future. The proceeds from these sales are held as liabilities and lead to the value

---

[2] Jay Ritter's website is: https://site.warrington.ufl.edu/ritter/ipo-data/

[3] The consumer sentiment index website is: http://www.sca.isr.umich.edu/

[4] Jeffery Wurgler's website is: https://pages.stern.nyu.edu/~jwurgler/

[5] The OECD database website is: https://data.oecd.org/

[6] The Economic Policy Uncertainty website is: https://www.policyuncertainty.com/index.html

become negative.

**Table 3–1. Summary statistics**

| | *Ps* | *Pc* | *CCI* | *GO* | *BW* | *NIPO* | *RIPO* | *VIX* | *IR* | *UEMP* | *GPR* | *CSI* |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Mean | 83.508 | 1.954 | 87.207 | 12.512 | -0.199 | 59.911 | 16.869 | 20.446 | 1.0213 | 6.485 | 2.630 | 82.055 |
| Median | 91.470 | 0.810 | 90.517 | 10.500 | -0.229 | 46.500 | 16.033 | 17.642 | 0.350 | 6.017 | 2.542 | 81.250 |
| Maximum | 269.000 | 55.660 | 133.900 | 64.333 | 1.099 | 401.000 | 47.133 | 51.910 | 5.423 | 13.033 | 3.965 | 98.933 |
| Minimum | 1.440 | -87.780 | 29.567 | 3.000 | -0.888 | 1.000 | -5.500 | 10.120 | 0.100 | 3.533 | 1.935 | 57.667 |
| Sad. Dev | 18.370 | 5.467 | 28.730 | 8.912 | 0.331 | 58.628 | 10.59 | 8.180 | 1.113 | 1.222 | 0.400 | 12.113 |
| Num of obs | 5320 | 5320 | 5320 | 5320 | 5320 | 5320 | 5320 | 5320 | 5320 | 5320 | 5320 | 5320 |
| Cross-sections | 95 | 95 | 95 | 95 | 95 | 95 | 95 | 95 | 95 | 95 | 95 | 95 |

* Ps and Pc refer to the percentage of capital amount of mutual funds invested in common stock, and cash, respectively. CCI means the consumer confidence index, GO means google trend searching volume of financial crisis and economic recession, BW indicates Baker and Wurgler index, NIPO means the number of IPOs, RIPO means the first day return of IPOs, VIX means the volatility index, IR means short-term interest rate, UEMP indicates the unemployment rate, and GPR refers to the geopolitical risk index. CSI indicates the consumer sentiment index.

### 3.2.2 Methods

Baker and Wurgler (2007)define 'investor sentiment' as investors' beliefs about cash flow and investment risk in the future. Therefore, it is easy to understand that the power or influence of investor sentiment cannot be detected by the market or investors immediately, which indicates that the market requires time to react to sentiment. The investors have limited ability to observe the movement of the economy, and most obtain delayed information, such as from the newspaper or internet. As a result, investors obtain information and adjust their behaviour gradually; for example, they may withdraw money from the stock market when pessimistic or invest more when optimistic. The market requires time to react to a change in the level of investor sentiment, indicating that sentiment is not an immediately effective indicator. Sentiment carries information into the future; its effect is expected to be delayed, implying that research should include lags in the model.

This research uses the panel autoregressive distributed lag model (ARDL), an extension based on the ARDL model proposed by Shin and Pesaran (1999), Pesaran et al. (1999), and Pesaran et al. (2001). This method includes appropriate amounts of lags for variables to analyse the long-term and short-term relationship or dynamic effect between portfolio composition and different investor sentiment measures. Additionally, ARDL can accommodate mixed integration variables; it can include both I(0) and I(1) into its estimation. It is unnecessary to engage a unique integration level because the ARDL model estimates appropriate cointegration for short-run and long-run coefficients. In addition, this approach incorporates lagged dependent variables and error correction terms, thereby substantially mitigating the issues of endogeneity and serial correlation(Zhang et al., 2021). Moreover, the panel ARDL allow long-run to homogeneous and heterogeneous in the short-run. This means that the coefficients representing the long-run effects of independent variables on the dependent variable are the same for all units. Conversely, the short-run dynamics (e.g., adjustments toward the long-run equilibrium) are allowed to be heterogeneous across units. This means that each unit can have its own unique short-run adjustment process, reflecting idiosyncratic factors like policy responses, local shocks, or differing adjustment speeds. The short-run heterogeneity allows the model to capture the diversity of behaviours and responses in the short term. According to D'Arcangelis and Rotundo (2021), mutual fund managers can be considered a homogeneous group due to their similar educational and professional backgrounds. Additionally, Mutual funds may also demonstrate herding behaviour, a phenomenon in finance where investors, including mutual fund managers, tend to follow the actions of others. This behaviour often results in similar investment patterns and portfolio compositions across funds, leading to a degree of homogeneity in their portfolios. Wermers (1999) documented that mutual funds have herding activities by analysing their portfolio from 1975-1994. Additionally, Santi and Zwinkels (2023)also illustrate the herding activity of mutual funds. The herding activities tend to increase after periods of significant market volatility and reduce with investor sentiment. Fund managers can also herd as a result of analysing of the same sentiment-related proxies(Liao et al., 2011).

The panel ARDL model (p, q, q, ..., q) of this study is as follows: p is the lag of the dependent variable, and q is the lag of the independent variables. Equation (3-1) can be restructured as an error correction model, Equation (3-2):

$$Pa_{j,t} = \alpha_0 + \beta_{1,j} Pa_{j,t-1} + \beta_2 S_{t-1} + \beta_3 IR_{t-1} + \beta_4 UEMP_{t-1} + \beta_5 GPR_{t-1} + \sum_{i=1}^{p-1} \lambda_{1,j}\, \Delta Pa_{j,t-i} + \sum_{i=0}^{q-1} \lambda_2 \Delta S_{t-i} + \sum_{i=0}^{q-1} \lambda_3\, \Delta IR_{t-i} + \sum_{i=0}^{q-1} \lambda_4\, \Delta UEMP_{t-i} + \sum_{i=0}^{q-1} \lambda_5\, \Delta GPR_{t-i} + \mu_t + \varepsilon_{j,t} \tag{3-1}$$

$$\Delta Pa_{j,t} = \phi_j ECT_{j,t-1} + \sum_{i=1}^{p-1} \lambda_{1,j}\, \Delta Pa_{j,t-i} + \sum_{i=0}^{q-1} \lambda_2 \Delta S_{t-i} + \sum_{i=0}^{q-1} \lambda_3\, \Delta IR_{t-i} + \sum_{i=0}^{q-1} \lambda_4\, \Delta UEMP_{t-i} + \sum_{i=0}^{q-1} \lambda_5\, \Delta GPR_{t-i} + \mu_t + \varepsilon_{j,t} \tag{3-2}$$

where the $Pa_t$ indicates one of two dependent variables $Ps_t$ or $Pc_t$, the percentage of capital mutual funds invest in common stock and cash. *S* represents arbitrary sentiment indicators in separate regressions. *CCI* is consumer confidence index; *NIPO* is the number of IPOs; *RIPO* is the first day return of IPOs; *GO* is the Google Trends search volume of financial crisis; *VIX* is the volatility index; *IR* is short-term interest rate; $UEMP$ is unemployment rate, and $GPR$ is the geopolitical risk index. $\Delta$ is the first-difference operator; $i = 0,1,2,3, \ldots, n$, $t = 1,2,3, \ldots., t$; $\varepsilon_t$ is the normal error term, and $ECT$ is the error correction term.

However, the CRSP database only provides mutual fund information quarterly, while the frequency of other independent variables is monthly. As a result, barring the number of IPOs, which is the sum of three months' value, to obtain quarterly data, the rest of the variables are transformed to quarterly data based on the three-month average value.

# 3.3 Results and Discussion

## 3.3.1 Panel Unit Root Test

Before exploring the relationship between sentiment and the percentage of capital invested in the common stock market and cash, it is necessary to conduct a unit root test to confirm the level of integration of each variable. Although panel ARDL can be used regardless of whether the variables are in order I(0) or I(1), the unit root test must still be used to detect whether I(2) or higher order variables appear because panel ARDL cannot be used when the integration order is greater than I(1).

However, it is necessary to test the cross-sectional dependence so at to decide which unit root test should be performed at first. I conduct cross-sectional dependence test proposed by Pesaran (2015) to independent variables. The null hypothesis is that there is no cross-sectional dependence. The results are presented in Table 3–2, which indicates that both Ps and Pc have cross-sectional dependence and need to use the second-generation unit root test(Pesaran, 2007). The null hypothesis is that the unit root presents for the variable. The unit root test results are also demonstrated in Table 3–3, which reveal that all the dependent variables are stationary at levels.

As for the independent variables, to test for a unit root (or stationarity), I use various tests in the panel dataset. Considering sample size and features, the Im, Pesaran, and Shin test proposed by Im et al. (2003) and the Fisher-ADF and Fisher-PP tests (Maddala and Wu, 1999) are used with lags selected by the Schwarz information criterion. For example, the Levin, Lin, and Chu test is suitable for the panels in which N < T, which is not the case in the current research. The applied tests have the null hypothesis that the variables have a unit root. Table 3–3 illustrates the results of different unit root tests in each series.

### Table 3–2. Cross sectional dependence test and CS-panel unit root test

| | *Ps* | | *Pc* | |
|---|---|---|---|---|
| **Pesaran CD** | 60.519* | | 50.180* | |
| **CIPS** | t-statistic | -4.356* | t-statistic | -2.839* |

Notes: * denotes statistically significant at 5% level. Ps and Pc refer to the percentage of capital amount of mutual funds invested in common stock, and cash, respectively.

### Table 3–3. Panel unit root test results

| Method Variables | *IPS* | *Δ* | *Fisher-ADF* | *Δ* | *Fisher-PP* | *Δ* |
|---|---|---|---|---|---|---|
| *CCI* | 5.684<br>[1.000] | -53.793*<br>[0.000] | 56.851<br>[1.000] | 2565.03*<br>[0.000] | 76.291<br>[1.000] | 2546.98*<br>[0.000] |
| *CSI* | -1.715*<br>[0.043] | -50.326*<br>[0.000] | 156.981<br>[0.962] | 2396.41*<br>[0.000] | 159.156<br>[0.950] | 3369.19*<br>[0.000] |
| *GO* | -22.401*<br>[0.000] | | 892.759*<br>[0.000] | | 938.280*<br>[0.000] | |
| *BW* | 16.285<br>[1.000] | -46.975*<br>[0.000] | 10.334<br>[1.000] | 2170.82*<br>[0.000] | 25.120<br>[1.000] | 2166.85*<br>[0.000] |
| *NIPO* | -4.914*<br>[0.000] | -29.867*<br>[0.000] | 221.777<br>[0.057] | 1268.82*<br>[0.000] | 592.489*<br>[0.000] | 4344.31*<br>[0.000] |
| *RIPO* | -21.345*<br>[0.000] | | 808.791*<br>[0.000] | | 793.702*<br>[0.000] | |
| *VIX* | -20.830*<br>[0.000] | -63.843*<br>[0.000] | 200.882<br>[0.280] | 3163.79*<br>[0.000] | 742.590*<br>[0.000] | 4787.20*<br>[0.000] |
| *IR* | -16.162*<br>[0.000] | | 232.220*<br>[0.020] | | 1086.10*<br>[0.000] | |
| *UEMP* | -11.738*<br>[0.000] | -80.546*<br>[0.000] | 165.871<br>[0.896] | 3962.84*<br>[0.000] | 126.503<br>[1.000] | 4231.20*<br>[0.000] |
| *GPR* | -32.631*<br>[0.000] | -116.651*<br>[0.000] | 166.063<br>[0.894] | 4900.91*<br>[0.000] | 160.430<br>[0.942] | 4037.16*<br>[0.000] |

Notes: * denotes statistically significant at 5% level. The results demonstrate coefficient and p-value with square brackets. Ps and Pc refer to the percentage of capital amount of mutual funds invested in common stock, and cash, respectively. CCI means the consumer confidence index, GO means google trend searching volume of financial crisis and economic recession, BW indicates Baker and Wurgler index, NIPO means the number of IPOs, RIPO means the first day return of IPOs, VIX means the volatility index, IR means short-term interest rate, UEMP indicates the unemployment rate, and GPR refers to the geopolitical risk index. *Δ* refers to the unit root test in the first difference.

### Table 3–4. Kao cointegration test results

| | ***Ps*** | ***Pc*** |
|---|---|---|
| | *ADF t-Statistic* | *ADF t-Statistic* |
| ***CCI*** | -10.284* | -15.451* |
| ***GO*** | -6.378* | 2.219* |
| ***BW*** | -5.830* | -5.229* |
| ***NIPO*** | -5.788* | -2.511* |
| ***RIPO*** | 3.917* | 3.854* |
| ***VIX*** | -6.521* | -4.683* |

Notes: * denotes statistically significant at 5% level.

As Table 3–3 above shows, the series are stationary except for *CCI*, *BW*, *NIPO*, *VIX*, *UEMP*, and *GPR*, which are I(1). Therefore, this dataset has a mixed order of integration, with no variable higher than I(1), which supports that panel ARDL is an appropriate method for this study.

The next step is to test the cointegration among all variables in each regression. This chapter employs Kao (1999) cointegration test. Table 3–4 demonstrates the result of cointegration test, it indicates that there all long-run relationship between variables since the null hypothesis of no cointegration is rejected.

### 3.3.2 Results of Panel ARDL Regressions

Before analysing the regression results, some terms need to be clarified. First, the long-run and short-run equations reflect the long-term and short-term influences, respectively. For example, a positive (negative) and statistically significant coefficient of an explanatory variable in the long-run (short-run) equation at the significance level of 1% shows that the explanatory variable is associated with a long-term (short-term) increase (decrease) of the dependent variable. It is an equilibrium path where short-run effects are adjusted to long-run equilibrium.

Second, COINTEQ represents an error correction term that implies the speed of adjustment from short-run equilibrium toward long-run equilibrium. Additionally, the error correction term should be negative, below one, and significant at the 5% level, which is fundamental to the necessity of efficiency and consistency of the long-run relationship among the variables in regressions. Eventually, this study selects the number of lags using the Schwarz information criteria. The significance level employed in this study is 5%.

## Table 3–5. The regression results of panel ARDL with Ps

| *Variables* | *Ps*<br>**(a)**<br>***Survey* base** | **(b)**<br>***Internet base*** | **(c)**<br>***Composite*** | **(d)**<br>***Market base*** | | |
|---|---|---|---|---|---|---|
| | | | ***Long run equation*** | | | |
| ***CCI*** | 0.041*<br>(0.005) | | | | | |
| ***GO*** | | -0.035*<br>(0.011) | | | | |
| ***BW*** | | | 0.613*<br>(0.305) | | | |
| ***NIPO*** | | | | 0.011*<br>(0.002) | | |
| ***RIPO*** | | | | | 0.073*<br>(0.010) | |
| ***VIX*** | | | | | | 0.039*<br>(0.016) |
| ***IR*** | 0.283*<br>(0.105) | 0.094<br>(0.113) | 0.226*<br>(0.102) | 0.371*<br>(0.101) | 0.391*<br>(0.052) | 0.125<br>(0.102) |
| ***UEMP*** | 0.375*<br>(0.078) | -0.051<br>(0.056) | 0.012<br>(0.058) | 0.006<br>(0.057) | -0.001<br>(0.055) | -0.088<br>(0.061) |
| ***GPR*** | -0.075<br>(0.242) | -0.298<br>(0.653) | 0.068<br>(0.698) | 0.531<br>(0.286) | 0.424<br>(0.274) | -0.035<br>(0.026) |
| | | | ***Short run equation*** | | | |
| ***COINTEQ*** | -0.352*<br>(0.024) | -0.349*<br>(0.023) | -0.352*<br>(0.022) | -0.350*<br>(0.023) | -0.350*<br>(0.023) | -0.345*<br>(0.022) |
| ***D(CCI)*** | -0.001<br>(0.008) | | | | | |
| ***D(GO)*** | | 0.008<br>(0.013) | | | | |
| ***D(BW)*** | | | 0.374<br>(0.417) | | | |
| ***D(NIPO)*** | | | | -0.001<br>(0.001) | | |
| ***D(RIPO)*** | | | | | -0.026*<br>(0.008) | |
| ***D(VIX)*** | | | | | | -0.020<br>(0.011) |
| ***D(IR)*** | -1.588*<br>(0.574) | -1.540*<br>(0.606) | -1.397*<br>(0.563) | -1.425*<br>(0.561) | -1.386*<br>(0.575) | -1.270*<br>(0.569) |
| ***D(UEMP)*** | -0.207*<br>(0.089) | -0.132<br>(0.092) | -0.125<br>(0.081) | -0.113<br>(0.090) | -0.102<br>(0.082) | -0.129<br>(0.087) |
| ***D(GPR)*** | 0.035<br>(0.337) | 0.066<br>(0.133) | -0.017<br>(0.125) | -0.079<br>(0.125) | -0.094<br>(0.127) | 0.003<br>(0.126) |
| ***C*** | 27.885*<br>(2.202) | 30.123*<br>(2.263) | 29.853*<br>(2. 178) | 28.930*<br>(2.179) | 28.816*<br>(2.172) | 29.263*<br>(2.161) |

Notes: * denotes statistically significant at 5% level. The results demonstrate coefficient and standard error with parentheses. Ps and Pc refer to the percentage of capital amount of mutual funds invested in common stock, and cash, respectively. CCI means the consumer confidence index, GO means google trend searching volume of financial crisis and economic recession, BW indicates Baker and Wurgler index, NIPO means the number of IPOs, RIPO means the first day return of IPOs, VIX means the volatility index, IR means short-term interest rate, UEMP indicates the unemployment rate, and GPR refers to the geopolitical risk index.

# Table 3–6. The regression results of panel ARDL with Pc

| *Variables* | *Pc* | | | | | |
|---|---|---|---|---|---|---|
| | **(a)**<br>***Survey* base** | **(b)**<br>***Internet base*** | **(c)**<br>***Composite*** | **(d)**<br>***Market base*** | | |
| | | | ***Long run equation*** | | | |
| ***CCI*** | -0.002*<br>(0.001) | | | | | |
| ***GO*** | | 0.004*<br>(0.001) | | | | |
| ***BW*** | | | 0.060<br>(0.033) | | | |
| ***NIPO*** | | | | -0.001<br>(0.001) | | |
| ***RIPO*** | | | | | -0.002*<br>(0.001) | |
| ***VIX*** | | | | | | 0.004*<br>(0.002) |
| ***IR*** | 0.025*<br>(0.011) | 0.038*<br>(0.010) | 0.044*<br>(0.011) | 0.042*<br>(0.012) | 0.035*<br>(0.013) | 0.038*<br>(0.011) |
| ***UEMP*** | -0.022*<br>(0.009) | 0.002<br>(0.005) | 0.013<br>(0.006) | 0.009<br>(0.006) | 0.005<br>(0.006) | 0.002<br>(0.006) |
| ***GPR*** | -0.007<br>(0.027) | 0.021<br>(0.023) | 0.008<br>(0.028) | 0.002<br>(0.032) | -0.012<br>(0.030) | 0.026<br>(0.028) |
| | | | ***Short run equation*** | | | |
| ***\|COINTEQ*** | 0.586*<br>(0.028) | -0.593*<br>(0.027) | -0.592*<br>(0.027) | -0.577*<br>(0.027) | -0.580*<br>(0.028) | -0.582*<br>(0.027) |
| ***D(CCI)*** | -0.022*<br>(0.009) | | -0.908*<br>(0.260) | | | |
| ***D(GO)*** | | 0.051*<br>(0.015) | | | | |
| ***D(BW)*** | | | -0.908*<br>(0.260) | | | |
| ***D(NIPO)*** | | | | -0.001<br>(0.001) | | |
| ***D(RIPO)*** | | | | | -0.012*<br>(0.005) | |
| ***D(VIX)*** | | | | | | 0.032*<br>(0.012) |
| ***D(IR)*** | -0.184<br>(0.234) | -0.920*<br>(0.175) | -0.444*<br>(0.197) | -0.407*<br>(0.200) | -0.447*<br>(0.199) | -0.502*<br>(0.187) |
| ***D(UEMP)*** | -0.138*<br>(0.039) | -0.119*<br>(0.029) | -0.082*<br>(0.027) | -0.061*<br>(0.029) | -0.046<br>(0.029) | -0.061*<br>(0.028) |
| ***D(GPR)*** | 0.405*<br>(0.107) | 0.323*<br>(0.099) | 0.200*<br>(0.091) | 0.166<br>(0.095) | 0.159<br>(0.095) | 0.176*<br>(0.092) |
| ***C*** | 1.156*<br>(0.112) | 0.803*<br>(0.110) | 0.870*<br>(0.111) | 0.848*<br>(0.108) | 0.904*<br>(0.109) | 0.792*<br>(0.107) |

Notes: * denotes statistically significant at 5% level. The results demonstrate coefficient and standard error with parentheses. Ps and Pc refer to the percentage of capital amount of mutual funds invested in common stock, and cash, respectively. CCI means the consumer confidence index, GO means google trend searching volume of financial crisis and economic recession, BW indicates Baker and Wurgler index, NIPO means the number of IPOs, RIPO means the first day return of IPOs, VIX means the volatility index, IR means short-term interest rate, UEMP indicates the unemployment rate, and GPR refers to the geopolitical risk index.

The regression results of all investor sentiment proxies are presented in Table 3–3 and Table 3–4. In the long run, the consumer confidence index has a significant relationship both types of holdings in the mutual funds. It shows a positive relationship between *CCI,* and the percentage of capital invested in the common stock market, whereas there is a negative relationship between *CCI* and the percentage of cash holding in the long run. Nevertheless, *CCI* only has a negative relationship with cash holdings in the short run. This indicates that mutual funds will invest more or change their portfolio allocation by deducting capital from cash and raising the common stock market investment when the consumer confidence index is at a higher level in the long run; the short-run relationship is insignificant.

The internet-based sentiment index, Google Trends search volume, has a significant negative relationship with the percentage of capital invested in the common stock market and a slight positive relationship with mutual fund cash holdings in the long-run. In the short-run, the internet-based sentiment index has an insignificant impact on the stock market investment. The higher the search volume, the lower the sentiment level. A higher search volume indicates that investors' concern about the financial crisis and economic recession is growing. This also implies that mutual funds will change their portfolio composition, reduce money in the common stock market when sentiment decreases, and raise safe assets to protect their position when facing short-term sentiment shocks.

The widely-known composite sentiment index (Baker and Wurgler, 2006; 2007) has been used in many studies, such as those by Corredor et al. (2013), Fong and Toh (2014), Huang et al. (2015), and Sibley et al. (2016). As a comprehensive market-based index regarded by many scholars, the BW Index has a correlation with portfolio composition; it only shows a positive correlation with the capital in stocks in the long run, which suggests that mutual funds will plough more capital into the stock market when sentiment is high. Nevertheless, the BW Index has limited crucial influence on the cash holdings in the long run. This result is consistent with several period studies, including those by Stambaugh et al. (2012) andBekiros et al. (2016), which imply that the BW Index has drawbacks and cannot cover some anomalies. However, in the short run, the BW Index suggests that mutual funds might reduce safe assets to chase more profits.

The market-based sentiment index, such as the number of IPOs issued (*NIPO*), only has a significant positive influence on the stock investment amount, which suggests that when more IPOs are issued, the mutual funds will put more capital into the stock market. This sentiment index has a negative impact on the percentage of cash holdings, which indicates that the sentiment will impact the mutual fund portfolio construction. Meanwhile, the first day return of IPOs (*RIPO*) displays a significant positive relationship to the capital amount of mutual funds invested in the stock market and a negative relationship with the amount of cash holdings. Thus, this still shows that with a higher investor sentiment index, mutual funds invest more in the stock market in the long-run, and sentiment impacts their portfolio composition. In addition, the results of *NIPO* and *RIPO* suggest that institutional investors may not follow the number of initial return rate of IPOs blindly; they may have their own standards for selecting IPOs, which is consistent with the study byNeupane et al. (2016). However, they still carry some information, which cannot be neglected.

As for the volatility index (*VIX*), Table 3–3 and Table 3–4 display the relationship between *VIX* and mutual fund portfolio composition. The results suggest that *VIX* has a significant positive relationship with mutual funds' stock investment at 5% level. Since the volatility index is also a metric for assessing price swings, mutual funds will have a greater opportunity to generate profits during periods of rapid price fluctuations; nevertheless, this comes with an increased level of risk. In addition, a higher *VIX* may present buying opportunities due to the lower price and exaggerate mutual funds' risk-seeking activities to obtain more profits. Additionally, *VIX* has significant positive impact on the mutual funds' cash holdings. This indicates mutual funds may increase their safe assets to avoid potential risk in the meantime.

Moreover, the interest rate and unemployment rate have a significant positive impact on capital in the stock market in the long run except for in *GO* and *VIX* regressions, but *IR* will influence stock investment negatively in the short run of all regressions. In addition, *IR* will have long run positive correlation with cash holdings in all the regressions. Ciminelli et al. (2022)document that higher interest rates are often a response to strong economic growth or expectations of such growth, and mutual funds may enhance their stock holdings to seize the anticipated market growth.

The unemployment rate only has a positive long-run relationship and negative short-run relationship with stock investment amount within the *CCI* regression. Additionally, *UEMP* only has a long-run negative impact on mutual funds cash holdings amount within *CCI* regression; however, it has a significantly negative influence on the cash holdings amount in the short run, except in *RIPO* regression. This indicates that mutual funds may consider that the profits of stock investment may exceed the interest income, but they may also increase their cash holding in the long run. Mutual funds may react to the short-term interest rate increase as a risk signal and reduce their stock exposure. When the unemployment rate is rising, the market may stay at a relatively low-price level, which might be a proper time to reduce cash holdings or invest other assets. This phenomenon could be attributed to the long-term investment strategies employed by institutional investors, such as mutual funds, which may lead them to enter the stock market at lower prices in response to rising unemployment rates. The result can be explained by the study of Atanasov (2021), which indicates that an increase in the unemployment rate, relative to its trend, is correlated with higher excess returns. Hence, it seems that investors perceive rising unemployment rates as opportunities for investment, likely driven by the anticipation of greater returns in these periods.

Finally, *GPR* has no significant long-run or short-run relationship with mutual funds stock investment percentage among all sentiment indicators; however, it possesses a positive impact on cash holdings in the short run with *CCI*, *GO*, and *VIX* regression. This indicates that mutual funds increase their cash holdings amount when the geopolitical risk index increases to avoid systematic risk.

In summary, all the investor sentiment proxies have a positively significant relationship with mutual funds' stock investment amount, except for *GO*, which has a negative impact. As for cash holdings, *CCI* and *RIPO* have a negatively significant correlation. *GO* and *VIX* have a significantly positive influence. However, *NIPO* and *BW* have no significant correlation with a mutual funds' cash amount. In addition, the composite BW Index has the highest correlation with both stock investments but no correlation with cash holdings. This implies that the BW Index only reflects partial information regarding a mutual fund's portfolio composition.

## 3.4 Robustness Test

Two procedures comprise the robustness test: conducting sub-panel regression to test robustness and applying Wald test. Next, I employ the consumer sentiment index, a widely used sentiment proxy, as a replacement index to confirm the robustness of relationship between investor sentiment and mutual fund portfolio composition. Finally, the Dumitrescu-Hurlin causality test is applied to investigate the causality among dependent variables and independent variables.

The sample period includes both the great financial crisis of 2007-2008 and the COVID-19 pandemic recession. Hence, I select the period which excludes serious crises and investor sentiment fluctuating rapidly, 2010 Q1 to 2019 Q4, to test whether sentiment proxies' impact mutual funds portfolio composition during a stable time.

# Table 3–7. Robustness check with sub-sample period of Ps

| *Variables* | *Ps* | | | | | |
|---|---|---|---|---|---|---|
| | **(a)** *Survey* **base** | **(b)** *Internet base* | **(c)** *Composite* | | **(d)** *Market base* | |
| | | | *Long run equation* | | | |
| ***CCI*** | 0.053* (0.016) | | | | | |
| ***GO*** | | -0.074 (0.042) | | | | |
| ***BW*** | | | 1.217* (0.610) | | | |
| ***NIPO*** | | | | 0.029* (0.005) | | |
| ***RIPO*** | | | | | 0.008 (0.013) | |
| ***VIX*** | | | | | | 0.001 (0.022) |
| ***IR*** | -0.313 (0.213) | 0.298 (0.105) | 0.277 (0.165) | 0.281* (0.147) | 0.207 (0.155) | 0.199 (0.167) |
| ***UEMP*** | 0.372* (0.170) | -0.049 (0.010) | -0.116 (0.072) | -0.105 (0.058) | -0.177* (0.063) | -0.207* (0.079) |
| ***GPR*** | 0.194 (0.250) | 0.460 (0.242) | 0.517* (0.263) | 0.784* (0.252) | 0.356 (0.269) | 0.358 (0.254) |
| | | | *Short run equation* | | | |
| ***COINTEQ*** | -0.342* (0.027) | -0.342* (0.024) | -0.348* (0.028) | -0.340* (0.027) | -0.342* (0.027) | -0.340* (0.028) |
| ***D(CCI)*** | -0.014 (0.009) | | | | | |
| ***D(GO)*** | | 0.027 (0.020) | | | | |
| ***D(BW)*** | | | 0.111 (0.421) | | | |
| ***D(NIPO)*** | | | | -0.001 (0.003) | | |
| ***D(RIPO)*** | | | | | -0.003 (0.006) | |
| ***D(VIX)*** | | | | | | 0.001 (0.013) |
| ***D(IR)*** | -0.291 (0.574) | -0.283 (0.574) | -0.157 (0.439) | -0.115 (0.462) | -0.270 (0.452) | -0.311 (0.466) |
| ***D(UEMP)*** | 0.211 (0.370) | 0.398 (0.351) | 0.328* (0.367) | 0.605 (0.352) | 0.400 (0.353) | 0.382 (0.347) |
| ***D(GPR)*** | 0.016 (0.112) | -0.023 (0.115) | -0.076 (0.112) | -0.112 (0.120) | -0.038 (0.116) | -0.035 (0.114) |
| ***C*** | 27.247* (2.374) | 29.603* (2.570) | 30.035* (2.610) | 28.527* (2.528) | 29.763* (2.593) | 29.762 (2.631) |

Notes: * denotes statistically significant at 5% level. The results demonstrate coefficient and standard error with parentheses. Ps and Pc refer to the percentage of capital amount of mutual funds invested in common stock, and cash, respectively. CCI means the consumer confidence index, GO means google trend searching volume of financial crisis and economic recession, BW indicates Baker and Wurgler index, NIPO means the number of IPOs, RIPO means the first day return of IPOs, VIX means the volatility index, IR means short-term interest rate, UEMP indicates the unemployment rate, and GPR refers to the geopolitical risk index.

## Table 3–8. Robustness check with sub-sample period of Pc

| *Variables* | | | *Pc* | | | |
|---|---|---|---|---|---|---|
| | **(a)** *Survey* **base** | **(b)** *Internet base* | **(c)** *Composite* | **(d)** *Market base* | | |
| | | | ***Long run equation*** | | | |
| ***CCI*** | 0.002 | | | | | |
| | (0.001) | | | | | |
| ***GO*** | | -0.004 | | | | |
| | | (0.004) | | | | |
| ***BW*** | | | 0.072 | | | |
| | | | (0.055) | | | |
| ***NIPO*** | | | | 0.0003 | | |
| | | | | (0.001) | | |
| ***RIPO*** | | | | | 0.001 | |
| | | | | | (0.001) | |
| ***VIX*** | | | | | | -0.001 |
| | | | | | | (0.002) |
| ***IR*** | -0.009 | 0.021 | 0.024 | 0.012 | 0.016 | 0.019 |
| | (0.018) | (0.015) | (0.015) | (0.140) | (0.013) | (0.015) |
| ***UEMP*** | 0.318* | 0.015 | 0.014* | 0.009 | 0.009 | 0.011 |
| | (0.032) | (0.009) | (0.006) | (0.005) | (0.053) | (0.007) |
| ***GPR*** | 0.040 | 0.054 | 0.055* | 0.046 | 0.051 | 0.046* |
| | (0.021) | (0.024) | (0.024) | (0.024) | (0.024) | (0.023) |
| | | | ***Short run equation*** | | | |
| ***COINTEQ*** | -0.610* | -0.596* | -0.615* | -0.599* | -0.605* | -0.605* |
| | (0.029) | (0.029) | (0.029) | (0.028) | (0.029) | (0.029) |
| ***D(CCI)*** | -0.024* | | | | | |
| | (0.009) | | | | | |
| ***D(GO)*** | | -0.002 | | | | |
| | | (0.018) | | | | |
| ***D(BW)*** | | | -0.282 | | | |
| | | | (0.489) | | | |
| ***D(NIPO)*** | | | | 0.003 | | |
| | | | | (0.002) | | |
| ***D(RIPO)*** | | | | | 0.005 | |
| | | | | | (0.005) | |
| ***D(VIX)*** | | | | | | 0.026 |
| | | | | | | (0.013) |
| ***D(IR)*** | -0.952* | -1.016* | -1.074* | -0.982* | -0.951* | -1.321* |
| | (0.427) | (0.453) | (0.459) | (0.444) | (0.420) | (0.483) |
| ***D(UEMP)*** | -0.329 | 0.020 | 0.075 | 0.069 | -0.002 | -0.106 |
| | (0.370) | (0.358) | (0.347) | (0.340) | (0.345) | (0.326) |
| ***D(GPR)*** | 0.281* | 0.174 | 0.189* | 0.212 | 0.185* | 0.209* |
| | (0.113) | (0.104) | (0.096) | (0.094) | (0.094) | (0.098) |
| ***C*** | 0.587* | 0.766* | 0.803* | 0.787* | 0.759* | 0.784* |
| | (0.123) | (0.127) | (0.131) | (0.125) | (0.121) | (0.123) |

Notes: * denotes statistically significant at 5% level. The results demonstrate coefficient and standard error with parentheses. Ps and Pc refer to the percentage of capital amount of mutual funds invested in common stock, and cash, respectively. CCI means the consumer confidence index, GO means google trend searching volume of financial crisis and economic recession, BW indicates Baker and Wurgler index, NIPO means the number of IPOs, RIPO means the first day return of IPOs, VIX means the volatility index, IR means short-term interest rate, UEMP indicates the unemployment rate, and GPR refers to the geopolitical risk index.

## Table 3–9. Full sample and sub-sample results comparison

| Variables | ***Ps*** | ***Pc*** | ***Ps*** | ***Pc*** |
|---|---|---|---|---|
| | ***Full sample*** | | ***Sub-sample*** | |
| ***CCI*** | 0.041* | -0.002* | 0.053* | 0.002 |
| | (0.005) | (0.001) | (0.016) | (0.001) |
| ***GO*** | -0.035* | 0.004* | -0.074 | -0.004 |
| | (0.011) | (0.001) | (0.042) | (0.001) |
| ***BW*** | 0.613* | 0.060 | 1.217* | -0.072 |
| | (0.305) | (0.033) | (0.610) | (0.055) |
| ***NIPO*** | 0.011* | -0.001 | 0.029* | 0.001 |
| | (0.002) | (0.001) | (0.005) | (0.001) |
| ***RIPO*** | 0.073* | -0.002* | 0.008 | 0.001 |
| | (0.010) | (0.001) | (0.013) | (0.001) |
| ***VIX*** | 0.039* | 0.004* | 0.001 | -0.001 |
| | (0.016) | (0.002) | (0.022) | (0.002) |

In Table 3-5, Table 3–6, and Table 3–7, the difference between the full sample and sub-sample is that the Google Trends index, first day return of IPOs, and volatility index no longer have a significant influence on mutual funds' stock investment, whereas all sentiment indicators have no significant correlation with mutual funds' cash holdings during stable period.

The reason for the differences may be that in stable market conditions, mutual funds have already set their portfolio strategies and allocations. Sudden changes in IPO returns might not significantly alter their investment decisions unless they foresee a longer-term trend or shift. According to Che-Yahya et al. (2014), institutional investors favour long-term capital appreciation and stable sources of revenue over immediate income. In addition, during stable times, the volatility index may relatively flatten and indicate lower risk; therefore, mutual funds may not be sensitive to the volatility index. Nevertheless, as for the cash holdings of mutual funds, investor sentiment only possesses limited influence toward them in stable market conditions. It is understandable that institutional investors may not alter their safe asset allocation frequently, to reduce their exposure when the market is relatively stable.

**Table 3–10. Robustness check with replacement sentiment index**

| Variables | *Ps* | *Pc* |
|---|---|---|
| | ***Long run equation*** | |
| ***CSI*** | 0.088* | -0.006* |
| | (0.013) | (0.001) |
| ***IR*** | 0.345* | 0.018 |
| | (0.112) | (0.011) |
| ***UEMP*** | 0.310* | -0.026* |
| | (0.084) | (0.009) |
| ***GPR*** | -0.592* | -0.034 |
| | (0.247) | (0.025) |
| | ***Short run equation*** | |
| ***\|COINTEQ*** | -0.347* | -0.582* |
| | (0.024) | (0.028) |
| ***D(CSI)*** | -0.026 | -0.020 |
| | (0.016) | (0.012) |
| ***D(IR)*** | -1.569* | -0.349 |
| | (0.564) | (0.203) |
| ***D(UEMP)*** | -0.228* | -0.086* |
| | (0.101) | (0.043) |
| ***D(GPR)*** | 0.152 | 0.214* |
| | (0.123) | (0.090) |
| ***C*** | 26.738* | 1.263* |
| | (2.136) | (0.113) |

Notes: * denotes statistically significant at 5% level. The results demonstrate coefficient and standard error with parentheses. Ps and Pc refer to the percentage of capital amount of mutual funds invested in common stock, and cash, respectively. CSI means the consumer sentiment index, IR means short-term interest rate, UEMP indicates the unemployment rate, and GPR refers to the geopolitical risk index.

As Table 3–8 demonstrates, the replacement sentiment index, the consumer sentiment index, has a significant correlation with both categories of mutual funds portfolio composition. This further confirms that the investor sentiment index significantly impacts mutual funds' assets allocation.

**Table 3–11. Wald test**

| | ***Ps*** | | ***Pc*** | |
|---|---|---|---|---|
| | *F-statistic* | *Chi-square* | *F-statistic* | *Chi-square* |
| ***CCI*** | 24.525* | 98.099* | 8.438* | 35.754* |
| ***GO*** | 5.370* | 21.478* | 9.059* | 36.235* |
| ***BW*** | 3.836* | 15.345* | 5.566* | 22.263* |
| ***NIPO*** | 13.988* | 55.951* | 4.530* | 18.121* |
| ***RIPO*** | 18.671* | 74.686* | 5.760* | 23.041* |
| ***VIX*** | 3.295* | 13.181* | 6.147* | 24.589* |

Notes: * denotes statistically significant at 5% level. CCI means the consumer confidence index, GO means google trend searching volume of financial crisis and economic recession, BW indicates Baker and Wurgler index, NIPO means the number of IPOs, RIPO means the first day return of IPOs, VIX means the volatility index.

To test the robustness further, it is necessary to test the joint significance of all explanatory variables in the regression, which can reflect whether the variables selected in the model are significant or efficient in the regressions of different investor sentiment indicators. The null hypothesis is $H_0: \beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$ and implies that the coefficients of all independent variables are equal to 0; if the null hypothesis is rejected, it indicates that all explanatory variables are crucial to being included in the model. As Table 3–9 demonstrates, the results imply that all mutual fund stock proportion and cash proportion regressions reject the null hypothesis. This implies that all the variables are efficient and important.

**Table 3–12. Dumitrescu-Hurlin causality test**

| | ***Ps*** | | | ***Pc*** | |
|---|---|---|---|---|---|
| | Wald-statistics | p-value | | Wald-statistics | p-value |
| *BW does not homogeneously cause Ps* | 8.253 | 0.000 | *BW does not homogeneously cause Pc* | 5.686 | 0.000 |
| *CCI does not homogeneously cause Ps* | 11.605 | 0.000 | *CCI does not homogeneously cause Pc* | 7.539 | 0.000 |
| *GO does not homogeneously cause Ps* | 11.929 | 0.000 | *GO does not homogeneously cause Pc* | 7.623 | 0.000 |
| *NIPO does not homogeneously cause Ps* | 8.941 | 0.000 | *NIPO does not homogeneously cause Pc* | 6.995 | 0.000 |
| *RIPO does not homogeneously cause Ps* | 5.581 | 0.000 | *RIPO does not homogeneously cause Pc* | 4.997 | 0.014 |
| *VIX does not homogeneously cause Ps* | 5.727 | 0.000 | *VIX does not homogeneously cause Pc* | 5.753 | 0.000 |
| *Ps does not homogeneously cause BW* | 6.328 | 0.000 | *Pc does not homogeneously cause BW* | 5.897 | 0.000 |
| *Ps does not homogeneously cause CCI* | 10.952 | 0.000 | *Pc does not homogeneously cause CCI* | 10.073 | 0.000 |
| *Ps does not homogeneously cause GO* | 20.896 | 0.000 | *Pc does not homogeneously cause GO* | 15.432 | 0.000 |
| *Ps does not homogeneously cause NIPO* | 11.043 | 0.000 | *Pc does not homogeneously cause NIPO* | 9.828 | 0.000 |
| *Ps does not homogeneously cause RIPO* | 5.530 | 0.000 | *Pc does not homogeneously cause RIPO* | 6.000 | 0.000 |
| *Ps does not homogeneously cause VIX* | 6.496 | 0.000 | *Pc does not homogeneously cause VIX* | 5.501 | 0.000 |

Notes: Ps and Pc refer to the percentage of capital amount of mutual funds invested in common stock, and cash, respectively. CCI means the consumer confidence index, GO means google trend searching volume of financial crisis and economic recession, BW indicates Baker and Wurgler index, NIPO means the number of IPOs, RIPO means the first day return of IPOs, VIX means the volatility index.

The final check is the causality test constructed by Dumitrescu and Hurlin (2012). The Dumitrescu-Hurlin panel causality test is a statistical method used to assess the direction of causality between variables in panel data settings. This test is particularly useful when dealing with non-stationary panel data and is designed to provide insights into whether one time series can be used to predict another within a panel data framework. Consequently, in the subsequent discussion, I endeavour to consistently remind the reader that the causal effects estimated herein should be interpreted as suggestive rather than conclusive. The null hypothesis in the Dumitrescu-Hurlin test posits that there is no causality between the variables across all the individual units in the panel. Conducting panel causality tests adds depth to the analysis by establishing the direction of causality, thereby offering a more complete and robust understanding of the interrelationships among the variables under study. As Table 3–10 illustrates, the conclusions of causality show the Wald-statistics are significant and reject the null hypothesis. These results imply that all investor sentiment indicators have causality toward the stock and cash percentage of mutual fund portfolios. In addition, the results indicate that there is bidirectional causality between investor sentiment and mutual funds portfolio composition. These suggest that not only does investor sentiment influence mutual fund investment decisions, but the collective actions of mutual funds feed back into the market's sentiment. Investor sentiment can act as a leading indicator for mutual fund investment decisions. When sentiment is positive, mutual funds increase their stock investment percentage, anticipating higher returns. Conversely, negative sentiment leads to a reduction in stock investments as funds become more risk averse.

A high level of pessimistic sentiment motivates investors to divest their holdings from the fund, consequently driving the fund manager to sell to raise cash, which could trigger the feedback loop(Feldman, 2010). The feedback loop is that the actions of mutual funds can reinforce existing sentiment trends. If mutual funds are buying stocks in a rising market (positive sentiment), it may increase the stock price and further elevate the market sentiment. Conversely, if mutual funds start selling stocks in a declining market (negative sentiment), it may lead to a decrease in stock prices. The falling prices could worsen investor sentiment, leading to more selling, and exacerbate pessimistic sentiment.

## 3.5 Conclusion

This study has investigated the relationship between investor sentiment and mutual fund portfolio composition in US market, addressing a gap in the existing literature. By incorporating multiple sentiment indices—including a survey-based consumer confidence index, an internet-based index derived from Google Trends search volume for negative sentiment keywords, a market-based index representing the number and first-day return of newly issued IPOs, the CBOE volatility index, and the Baker and Wurgler Index—this research provides a comprehensive analysis of how investor sentiment influences professional portfolio management. Empirical results demonstrate that most investor sentiment indicators, including survey-based, market-based, and internet-based indicators, significantly influence mutual fund portfolio composition. This underscores the importance of sentiment in shaping the investment decisions of institutional investors. These finding highlights that not all sentiment measures are equally informative regarding portfolio composition. Considering the correlation of investor sentiment indicators, although the BW Index has the highest correlation with stock investments in a mutual fund's portfolio, compared to single investor sentiment indicators, it has no correlation with mutual funds' cash holding amounts. This implies that the BW Index only carries partial information concerning mutual funds portfolio composition. Therefore, it may be interesting to compose a new composite index to cover a more comprehensive scope of portfolio composition.

In sum, this study contributes to the literature by highlighting the nuanced relationship between investor sentiment and mutual fund portfolio composition. It emphasizes the need for more robust composite sentiment measures that can capture the multifaceted nature of institutional investment strategies.

# Chapter 4. Will a Portfolio of Mutual Funds React to Investor Sentiment Fluctuations Asymmetrically?

## 4.1 Introduction

The previous chapter elucidates the linear relationship between various investor sentiment indicators and mutual fund portfolio composition. This chapter determines whether there is a nonlinear asymmetric nexus among investor sentiment and portfolio composition. Previous studies have discovered that investor sentiment may have an asymmetrical influence on the market. For instance, Yu and Yuan (2011)demonstrate that the significant positive nexus between the aggregate market's expected return and its conditional volatility breaks down following periods of high sentiment. Stambaugh et al. (2015) find that sentiment-driven traders have a sharply positive demand on many stocks when sentiment is at a high level, whereas a corresponding strongly negative demand on stocks disappears when sentiment is low. This implies that investor sentiment will impact arbitrage activities asymmetrically. Therefore, trading in the stock market has an idiosyncratic impact during the declining and rising sentiment periods. In relation to market sentiment fluctuations, sentiment-driven investing is asymmetrical. This disparity is also apparent when comparing conduct in bear versus bull markets, as investors are more susceptible to sentiment-driven trading in bear markets(Chau et al., 2016). Jiang et al. (2018) show that numerous asymmetric risk measures are sensitive to investor sentiment.

Additionally, the stock return and volatility will react to sentiment asymmetrically in the Indian market(Chakraborty and Subramaniam, 2020). Thus, it is feasible to incorporate prospect theory into detecting the asymmetric effect of investor sentiment. As a result, numerous asymmetric anomalies resulting from periods of high and low sentimental influence are discovered. The question is whether investor sentiment will have an asymmetric impact on mutual fund portfolio compositions. Many existing studies indicate that investor sentiment has a relationship with mutual funds by using mutual fund flow to evaluate sentiment. Baker and Wurgler (2007) find that their principal component analysis (PCA) composite index which including 6 market-based proxies has a relationship with mutual fund flow, and it is a standard for valuing over- and underreaction. Furthermore, Ben-Rephael et al. (2012), Akbas et al. (2015), and DeVault et al. (2019) use mutual fund flow to gauge investor sentiment.

In the previous chapter, I constructed the linear ARDL model in the panel to evaluate the symmetric relationship between investor sentiment and mutual fund portfolio composition. Will investor sentiment generate an asymmetric effect on the portfolio composition, or will the increasing or decreasing sentiment have the same impact on the portfolio composition? This chapter determines whether an asymmetric effect exists between investor sentiment and portfolio composition, in which the portfolio composition may react differently to positive and negative sentiment shocks.

In exploring the asymmetric effect of investor sentiment on mutual fund portfolio composition, this research integrates the principles of prospect theory, a cornerstone of behavioural economics developed by Kahneman and Tversky (1979). Prospect theory, which challenges the traditional assumptions of rational choice under uncertainty, is particularly relevant in understanding investor behaviour in financial markets. It posits that people value gains and losses differently, leading to decision-making that deviates from classical utility theory.

The inclusion of prospect theory in this context is motivated by its potential to provide a more nuanced understanding of how investors' sentiment influences their portfolio choices. Specifically, the theory's value function, which reflects how investors perceive gains and losses relative to a specific reference point, offers a compelling framework to analyse the impact of sentiment-driven decisions on mutual fund compositions. The asymmetry in investor response to positive and negative market developments, as articulated by the value function, aligns with the observed patterns of sentiment-driven investment behaviours. The existing literature mainly focuses on the mutual fund flows under the view of prospect theory. Firstly, a definition needs to be introduced, the Prospect Theory (PT) values. It refers to a measure that evaluates the attractiveness of a portfolio's return distribution based on the principles of Prospect Theory. For example, Gu and Yoo (2021)find that funds with higher PT values attract greater net flows from investors, indicating a preference for portfolios that align with investors' risk preferences as described by Prospect Theory. Similarly, Wang and Han (2023)report a positive correlation between PT values and mutual fund flows in the Chinese market. Investors are assumed to be driven by their attitudes towards gains and losses, with a greater emphasis on avoiding losses than on maximizing gains, which known as loss aversion. In addition, Giannikos et al. (2023)highlight how institutional managers demonstrate "play it safe" behaviour consistent with Prospect Theory by favouring blind principal bids (BPBs) over agency trades. This preference arises from the desire to avoid potentially higher variable costs by opting for the fixed cost of BPBs.

Overall, by concentrating on the value function, this study aims to shed light on how investor sentiment, particularly in response to gains and losses, shapes mutual fund portfolio composition under prospect theory. This focus aligns with the broader objective of elucidating the behavioural underpinnings of investment decisions in the face of market uncertainties and highlighting the importance of understanding behavioural biases in portfolio construction in the financial industry. However, this study focuses solely on the value function aspect of prospect theory, while deliberately excluding the weighting function from its scope which could be regarded as a limitation compared to other research. The reasons for this exclusion are threefold. First, the primary methodology employed in this study is linear form regression. The weighting function of prospect theory, by contrast, is a nonlinear choice function. Including it introduces a level of complexity not compatible with the linear analytical framework of this research. Second, the weighting function in prospect theory relies heavily on choice probability data to interpret how individuals perceive the probability of outcomes. Unfortunately, access to such data is beyond the scope of this research, rendering a comprehensive analysis of the weighting function unfeasible. Last, the nature of the proposed model limits the extent to which it can comprehensively examine all constructs derived from prospect theory. Given these constraints, the research focuses on the more directly applicable aspects of the theory—namely, the reference point and loss aversion concepts.

This chapter proceeds in four sections. The next section outlines the data sample and methodology. Section 4.3 demonstrate the empirical results. Section 4.4 illustrate the robustness check. And the last section provides conclusion.

# 4.2 Data and Methods

The dataset of this chapter is the same as Chapter 3. In this section, I employ the nonlinear ARDL model (Shin et al., 2014) in panel form to detect the asymmetric effects and the cointegrating relationship between the relevant factors. By employing this model, I can distinguish between the long-term and short-term effects of explanatory variables on the dependent variable. Furthermore, the asymmetric connection may be clearly represented. The nonlinear panel ARDL is an extension based on the ARDL model proposed by Shin and Pesaran (1999) and Pesaran et al. (1999; Pesaran et al. (2001). This model does not have a convergence problem like other widely used asymmetric models, including the nonlinear threshold vector error correction model or the smooth transition model, due to the excessive number of parameters to be estimated. Moreover, the model offers several advantages, including its capacity to handle variables with different orders of integration (I(0) and I(1)), especially, the dataset in this thesis has mixed order of integration. In addition, its versatility in modelling both short-term and long-term effects, and its ability to address endogeneity issues.(Kassouri and Altıntaş, 2020; Suanin, 2021; Wang et al., 2021; Sebri et al., 2023). Therefore, considering the superior advantages of the nonlinear panel ARDL approach, many researchers have applied this method in recent years(Salisu and Isah, 2017; Ugurlu-Yildirim et al., 2021).

Thus, the equation of the long-run cointegration model according to the study by Shin et al. (2014) is presented as follows:

$$y_{j,t} = \beta_1 S_t^+ + \beta_2 S_t^- + \gamma c_t + \varepsilon_{j,t} \tag{4-1}$$

where $y_{j,t}$ refers to dependent variables $Pa_{j,t}$, which are the percentage of capital amount invested in common stock and cash of mutual fund $j$ over a period of $t$. $\beta^+$

and $\beta^-$ indicate the long run parameters. $S_t$ refers to arbitrary investor sentiment variables entering the model asymmetrically, which include *CCI* (consumer confidence index), *GO* (Google Trends), *BW* (BW Index), *NIPO* (number of IPOs), *RIPO* (the first day return of IPOs), and *VIX* (volatility index). $c_t$ refers to a (g×1) vector of control variables entering the model symmetrically(Shin et al., 2014), including *IR* (interest rate), *UEMP* (unemployment rate), and *GPR* (geopolitical risk index). $S_t$ can be decomposed as $S_t = S_0 + S_t^+ + S_t^-$ where $S_0$ is the initial value of each time point *t*, which could be regarded as the reference point; $S_t^+$and $S_t^-$ refer to the partial sum processes of positive and negative shocks in investor sentiment indicators separately(Samargandi, 2019):

$$S_t^+ = \sum_{k=1}^{t} \Delta S_k^+ = \sum_{k=1}^{t} max(\Delta S_k, 0) \tag{4-2}$$

$$S_t^- = \sum_{k=1}^{t} \Delta S_k^- = \sum_{k=1}^{t} min(\Delta S_k, 0) \tag{4-3}$$

where *k* refers to the sample unit, and *t* refers to the number of periods.

I extend Equation (4-1) to the nonlinear panel ARDL, Equation (4-4), by incorporating long-run and short-run asymmetry relationships in the linear panel ARDL.

$$\begin{aligned} Pa_{j,t} = {} & \alpha_0 + \beta_{0,j} Pa_{j,t-1} + \beta_1 S_{t-1}^+ + \beta_2 S_{t-1}^- + \beta_3 IR_{t-1} + \beta_4 UEMP_{t-1} + \beta_5 GPR_{t-1} + \\ & \sum_{i=1}^{p-1} \lambda_{0,j}\, \Delta Ps_{j,t-i} + \sum_{i=0}^{q-1} \lambda_1\, \Delta S_{t-i}^+ + \sum_{i=0}^{q-1} \lambda_2\, \Delta S_{t-i}^- + \sum_{i=0}^{q-1} \lambda_3\, \Delta IR_{t-i} + \sum_{i=0}^{q-1} \lambda_4\, \Delta UEMP_{t-i} + \\ & \sum_{i=0}^{q-1} \lambda_5\, \Delta GPR_{t-i} + \mu_t + \varepsilon_{j,t} \end{aligned} \tag{4-4}$$

where $\mu_t$ refers to group-specific effect.

As Pesaran et al. (2001) proposes, Equation (4-4) can be reparametrized as an error correction model; thus, Equation (4-5):

$$\Delta Pa_{j,t} = \phi_j ECT_{j,t-1} + \sum_{i=1}^{p-1} \lambda_{0,j}\, \Delta Pa_{j,t-i} + \sum_{i=0}^{q-1} \lambda_1\, \Delta S_{t-i}^{+} + \sum_{i=0}^{q-1} \lambda_2\, \Delta S_{t-i}^{-} + \sum_{i=0}^{q-1} \lambda_3\, \Delta IR_{t-i} + \sum_{i=0}^{q-1} \lambda_4\, \Delta UEMP_{t-i} + \sum_{i=0}^{q-1} \lambda_5\, \Delta GPR_{t-i} + \mu_t + \varepsilon_{j,t} \quad (4\text{-}5)$$

The error correction term represents the deviation from the long-term equilibrium in the data. Expressly, it signifies the speed at which a dependent variable returns to equilibrium after a change in a short-run effect of independent variable. This 'speed of adjustment' is a major feature in error correction models.

## 4.3 Results and Discussion

Before analysing the variable coefficients, it is crucial to test the unit root and cointegration at first. The panel unit root results are presented in Chapter 3, and I use the same dataset. The panel unit root test results indicate the dataset has a mixed order of integration. Thus, the nonlinear panel ARDL is a proper method for this study.

**Table 4–1. Kao cointegration test results**

| | ***Ps*** | ***Pc*** |
|---|---|---|
| | *ADF t-Statistic* | *ADF t-Statistic* |
| ***CCI⁺ and CCI⁻*** | -14.182* | -13.751* |
| ***GO⁺ and GO⁻*** | -9.997* | 3.187* |
| ***BW⁺ and BW⁻*** | -9.195* | -4.818* |
| ***NIPO⁺ and NIPO⁻*** | -10.166* | -3.086* |
| ***RIPO⁺ and RIPO⁻*** | -7.753* | 3.854* |
| ***VIX⁺ and VIX⁻*** | -10.352* | -2.971* |

Notes: * denotes statistically significant at 5% level.

The next step is to perform panel cointegration test proposed by Kao (1999). As Table 4–1 shows, that all the variables in each regression have cointegration at 5% level. It suggests that all variables have long-run relationships with each other.

It is also important to note that the estimated coefficient for the error correction mechanism (COINTEQ) is significant in all cases, falling within the range of -1 to 0. This indicates that, on average, shocks and deviations from the long-run path are corrected.

## Table 4–2. Nonlinear ARDL regression result of Ps

| *Variables* | (a)<br>*Survey* base | (b)<br>*Internet base* | (c)<br>*Composite* | (d)<br>*Market base* | | |
|---|---|---|---|---|---|---|
| | | | *Ps* | | | |
| | | | *Long run equation* | | | |
| ***CCI*** $^{+}$ | 0.023* | | | | | |
| | (0.008) | | | | | |
| ***CCI*** $^{-}$ | 0.009 | | | | | |
| | (0.014) | | | | | |
| ***GO*** $^{+}$ | | -0.059* | | | | |
| | | (0.019) | | | | |
| ***GO*** $^{-}$ | | -0.086* | | | | |
| | | (0.019) | | | | |
| ***BW*** $^{+}$ | | | 0.054 | | | |
| | | | (0.282) | | | |
| ***BW*** $^{-}$ | | | -0.907* | | | |
| | | | (0.311) | | | |
| ***NIPO*** $^{+}$ | | | | 0.008* | | |
| | | | | (0.004) | | |
| ***NIPO*** $^{-}$ | | | | 0.004 | 0.020 | |
| | | | | (0.004) | (0.012) | |
| ***RIPO*** $^{+}$ | | | | | 0.008 | |
| | | | | | (0.014) | |
| ***RIPO*** $^{-}$ | | | | | | |
| ***VIX*** $^{+}$ | | | | | | 0.056* |
| | | | | | | (0.018) |
| ***VIX*** $^{-}$ | | | | | | 0.034 |
| | | | | | | (0.018) |
| ***IR*** | 0.406* | 0.298 | 0.210 | 0.403* | 0.297* | 0.345* |
| | (0.182) | (0.105) | (0.127) | (0.137) | (0.130) | (0.140) |
| ***UEMP*** | 0.300* | -0.049 | 0.080 | 0.156* | 0.145* | 0.088 |
| | (0.079) | (0.010) | (0.068) | (0.072) | (0.068) | (0.073) |
| ***GPR*** | 0.149 | 0.460 | 0.353 | 0.671* | 0.390 | 0.322 |
| | (0.292) | (0.242) | (0.262) | (0.297) | (0.268) | (0.263) |
| | | | *Short run equation* | | | |
| ***\|COINTEQ*** | -0.360* | -0.396* | -0.382* | -0.382* | -0.371* | -0.373* |
| | (0.024) | (0.027) | (0.025) | (0.024) | (0.025) | (0.025) |
| ***D(CCI*** $^{+}$***)*** | 0.008 | | | | | |
| | (0.011) | | | | | |
| ***D(CCI*** $^{-}$***)*** | 0.016 | | | | | |
| | (0.018) | | | | | |
| ***D(GO*** $^{+}$***)*** | | 0.035 | | | | |
| | | (0.019) | | | | |
| ***D(GO*** $^{-}$***)*** | | 0.028 | | | | |
| | | (0.049) | | | | |
| ***D(BW*** $^{+}$***)*** | | | 1.478* | | | |
| | | | (0.320) | | | |
| ***D(BW*** $^{-}$***)*** | | | -1.980 | | | |
| | | | (1.205) | | | |
| ***D(NIPO*** $^{+}$***)*** | | | | -0.003 | | |
| | | | | (0.002) | | |
| ***D(NIPO*** $^{-}$***)*** | | | | 0.001 | | |
| | | | | (0.002) | | |
| ***D(RIPO*** $^{+}$***)*** | | | | | -0.035 | |
| | | | | | (0.021) | |
| ***D(RIPO*** $^{-}$***)*** | | | | | -0.001 | |
| | | | | | (0.010) | |
| ***D(VIX*** $^{+}$***)*** | | | | | | -0.051* |
| | | | | | | (0.012) |
| ***D(VIX*** $^{-}$***)*** | | | | | | 0.062 |
| | | | | | | (0.034) |
| ***D(IR)*** | -1.703* | -2.157* | -1.709* | -1.621* | -1.674* | -1.587* |
| | (0.583) | (0.868) | (0.579) | (0.569) | (0.600) | (0.610) |
| ***D(UEMP)*** | -0.163 | -0.213 | -0.182* | -0.181* | -0.158* | -0.155* |
| | (0.096) | (0.108) | (0.082) | (0.092) | (0.077) | (0.083) |
| ***D(GPR)*** | -0.076 | 0.046 | -0.117 | -0.114 | -0.101 | 0.030 |
| | (0.121) | (0.134) | (0.385) | (0.124) | (0.132) | (0.131) |
| ***C*** | 29.231* | 32.195* | 30.760* | 30.643* | 30.083* | 30.495* |
| | (2.262) | (2.489) | (2.346) | (2.266) | (2.362) | (2.331) |

Notes: * denotes statistically significant at 5% level. The results demonstrate coefficient and standard error with parentheses. + indicates positive change, - indicates negative change, Ps and Pc refer to the percentage of capital amount of mutual funds invested in common stock, and cash, respectively. CCI means the consumer confidence index, GO means google trend searching volume of financial crisis and economic recession, BW indicates Baker and Wurgler index, NIPO means the number of IPOs, RIPO means the first day return of IPOs, VIX means the volatility index, IR means short-term interest rate, UEMP indicates the unemployment rate, and GPR refers to the geopolitical risk index.

# Table 4–3. Nonlinear ARDL regression result of Pc

| *Variables* | Pc | | | | | |
|---|---|---|---|---|---|---|
| | (a) *Survey* base | (b) *Internet base* | (c) *Composite* | (d) *Market base* | | |
| | | | ***Long run equation*** | | | |
| ***CCI*** $^{+}$ | -0.003* | | | | | |
| | (0.001) | | | | | |
| ***CCI*** $^{-}$ | -0.004* | | | | | |
| | (0.002) | | | | | |
| ***GO*** $^{+}$ | | -0.001 | | | | |
| | | (0.002) | | | | |
| ***GO*** $^{-}$ | | 0.001 | | | | |
| | | (0.002) | | | | |
| ***BW*** $^{+}$ | | | 0.091* | | | |
| | | | (0.032) | | | |
| ***BW*** $^{-}$ | | | 0.139* | | | |
| | | | (0.035) | | | |
| ***NIPO*** $^{+}$ | | | | 0.0006 | | |
| | | | | (0.001) | | |
| ***NIPO*** $^{-}$ | | | | 0.001* | | |
| | | | | (0.001) | | |
| ***RIPO*** $^{+}$ | | | | | 0.002 | |
| | | | | | (0.001) | |
| ***RIPO*** $^{-}$ | | | | | 0.003 | |
| | | | | | (0.002) | |
| ***VIX*** $^{+}$ | | | | | | 0.006* |
| | | | | | | (0.002) |
| ***VIX*** $^{-}$ | | | | | | 0.008* |
| | | | | | | (0.002) |
| ***IR*** | 0.056* | 0.006 | 0.048* | 0.064* | 0.050* | 0.018 |
| | (0.020) | (0.012) | (0.014) | (0.015) | (0.016) | (0.016) |
| ***UEMP*** | -0.017* | -0.013* | 0.008 | 0.005 | -0.002 | -0.016* |
| | (0.009) | (0.006) | (0.008) | (0.008) | (0.008) | (0.008) |
| ***GPR*** | 0.032 | -0.0005 | 0.038 | 0.043 | -0.018 | 0.006 |
| | (0.032) | (0.021) | (0.029) | (0.034) | (0.032) | (0.029) |
| | | | ***Short run equation*** | | | |
| ***\|COINTEQ*** | -0.585* | -0.638* | -0.614* | -0.607* | -0.585* | -0.594* |
| | (0.028) | (0.030) | (0.028) | (0.028) | (0.028) | (0.028) |
| ***D(CCI*** $^{+}$***)*** | -0.026* | | | | | |
| | (0.01) | | | | | |
| ***D(CCI*** $^{-}$***)*** | -0.046* | | | | | |
| | (0.020) | | | | | |
| ***D(GO*** $^{+}$***)*** | | 0.082* | | | | |
| | | (0.023) | | | | |
| ***D(GO*** $^{-}$***)*** | | -0.006 | | | | |
| | | (0.017) | | | | |
| ***D(BW*** $^{+}$***)*** | | | 0.535 | | | |
| | | | (0.316) | | | |
| ***D(BW*** $^{-}$***)*** | | | -3.586* | | | |
| | | | (0.921) | | | |
| ***D(NIPO*** $^{+}$***)*** | | | | -0.002 | | |
| | | | | (0.002) | | |
| ***D(NIPO*** $^{-}$***)*** | | | | -0.0004 | | |
| | | | | (0.001) | | |
| ***D(RIPO*** $^{+}$***)*** | | | | | -0.025* | |
| | | | | | (0.012) | |
| ***D(RIPO*** $^{-}$***)*** | | | | | -0.002 | |
| | | | | | (0.008) | |
| ***D(VIX*** $^{+}$***)*** | | | | | | 0.034* |
| | | | | | | (0.012) |
| ***D(VIX*** $^{-}$***)*** | | | | | | 0.019 |
| | | | | | | (0.023) |
| ***D(IR)*** | -0.129 | -0.723* | -0.433* | 0.434* | -0.447* | -0.444 |
| | (0.256) | (0.220) | (0.197) | (0.200) | (0.199) | (0.189) |
| ***D(UEMP)*** | -0.171* | -0.101* | -0.086* | -0.065* | -0.034* | -0.444 |
| | (0.050) | (0.030) | (0.027) | (0.031) | (0.030) | (0.189) |
| ***D(GPR)*** | 0.405* | 0.332* | 0.107 | 0.157 | 0.125 | 0.166 |
| | (0.109) | (0.097) | (0.102) | (0.094) | (0.103) | (0.090) |
| ***C*** | 0.806* | 0.998* | 0.693* | 0.900* | 1.014* | 1.000* |
| | (0.126) | (0.131) | (0.133) | (0.117) | (0.117) | (0.118) |

Notes: * denotes statistically significant at 5% level. The results demonstrate coefficient and standard error with parentheses. + indicates positive change, - indicates negative change, Ps and Pc refer to the percentage of capital amount of mutual funds invested in common stock, and cash, respectively. CCI means the consumer confidence index, GO means google trend searching volume of financial crisis and economic recession, BW indicates Baker and Wurgler index, NIPO means the number of IPOs, RIPO means the first day return of IPOs, VIX means the volatility index, IR means short-term interest rate, UEMP indicates the unemployment rate, and GPR refers to the geopolitical risk index.

As Table 4–2 and Table 4–3 demonstrate, $CCI^+$ and $CCI^-$ capture the positive and negative shocks of investor sentiment. The results indicate that positive shocks of *CCI* have significant impact on mutual funds stock investment proportion, whereas the significance of negative shocks is absent in the long run. As for cash holdings, both positive shocks and negative shocks have significant impact. The coefficient of $CCI^+$ is larger than $CCI^-$. This implies that mutual fund stock investments are more sensitive to positive shocks than negative shocks in the long run, and that fluctuations in investor sentiment increase mutual fund risky assets like stock investments more during periods of positive shocks of sentiment than those of negative shocks. This result is consistent with Fong (2013) conclusion that high sentiment may encourage risk-seeking behaviour to invest in the stock market. Higher sentiment may accompany higher risk because the stock price may drift from its fundamental value and lower the return.

However, the sentiment fluctuations may have a symmetric effect to cash holdings. Both the positive and negative shocks of investor sentiment have a statistically significant negative effect on cash holding percentage in the long run, which implies that the mutual funds decrease cash holdings during periods with increasing sentiment and increase cash holdings during those with decreasing sentiment. The results, however, show that *CCI* has no significant influence on portfolio composition in the short run toward risky assets but has a significant negative impact on cash proportion.

Additionally, Table 4–2 and Table 4–3 report the nonlinear panel ARDL estimation results of internet-based sentiment indicators, the Google Trends search volume (*GO*), and mutual fund portfolio composition. There are two coefficients,

$GO^+$ and $GO^-$, where both the positive and negative shocks are significantly correlated with the amount of stock investments in mutual funds. This shows how the fluctuation of the Google search volume of the terms 'financial crisis' and 'economic recession' affects the amount of stock investments in mutual funds in the long run. Mutual funds react to negative shocks more than positive shocks. A higher *GO* means concern of a crisis is increasing, which suggests potential risk or uncertainty exists. Therefore, mutual funds may reduce their stock investments to avoid risk. This suggests that mutual funds react to a low risk level more than a higher risk level, and they increase their stock holdings during *GO* decreasing periods more than they decrease their stock holdings during *GO* increasing periods. In the short run, neither the positive nor negative changes have significant influence on mutual funds' stock investments. Considering the cash amount of mutual funds, the coefficients do not show a significant impact on cash holdings from positive or negative shocks of *GO* in the long run. However, in the short run, the mutual fund will increase cash holdings during the positive change period, which implies that mutual funds will increase the safe asset amount when *GO* increases in the short run to prepare for the coming uncertainty.

Next, I discuss the relationship between the BW Index and mutual fund portfolio composition. In the long run, the results show that the BW Index has significant asymmetric effects on mutual funds' stock investment capital. The stock proportion is not impacted by positive shocks to the BW Index. However, the negative change in the BW Index has a significantly negative long-run effect on the capitalization percentage of mutual funds in the stock market, and they react less to positive changes than negative changes. There is also evidence to show that the BW Index influences mutual fund stock investments asymmetrically in the short run. The dependent variable only significantly correlates with to positive shocks, and there is no correlation with negative BW Index changes. Furthermore, considering the amount of cash holdings, both positive and negative shocks in the BW Index have a long-run significant relationship. However, in the short run, cash holdings only have a significantly negative relationship with negative changes in the BW Index. In addition, the coefficient of negative shocks is larger than that of positive shocks, which implies mutual funds' cash amounts may be more sensitive to a declining BW Index in both the long run and short run.

The coefficients $NIPO^{+}$ and $NIPO^{-}$ capture the significantly asymmetric effect of the number of IPOs (*NIPO*) on mutual funds' capital amounts in the stock market in the long run. The results imply that the positive change in *NIPO* will attract more investment from mutual funds into the stock market, while the negative changes in *NIPO* have no significant relationship with mutual funds' stock investment proportion in the long run. Nevertheless, the negative shocks of *NIPO* have a small but significant relationship with cash holdings of mutual funds, for which the coefficient is 0.001. That suggests the fluctuation of *NIPO* only has a very limited effect on the cash holdings of mutual funds. Additionally, in the short run, neither positive nor negative shocks have a significant nexus with the proportion of stocks or cash holdings owned by mutual funds. This indicates mutual funds will not react to the *NIPO* immediately, which may be due to mutual funds selecting IPOs based on their quality rather than investing blindly. According to Neupane et al. (2016), institutional investors like mutual funds alter their allocations differently for cold, warm, and hot IPOs, respectively.

As shown in Table 4–2 and Table 4–3, both the positive and negative shocks of first-day return rate of IPOs (*RIPO*) has no significant impact on the percentage of stock investment and cash holdings of mutual funds' portfolios in the long run. In the short run, neither positive nor negative shocks impact capital amount of stock investment. Nevertheless, positive *RIPO* shocks will have a significant negative impact on cash holding amounts, whereas negative changes have no significant impact. The results suggest that *RIPO* has limited power over finding asymmetric effects on mutual funds' portfolio composition; this indicates that mutual funds may only reduce their cash holdings during periods of *RIPO* positive shocks in the short term. This finding is consistent with Funaoka and Nishimura (2019) conclusion that institutional investors may have a private information advantage on IPOs' initial returns and set a planned favourable return for them. Therefore, this may explain why mutual fund portfolio compositions do not react well to positive or negative changes in *RIPO*.

The results demonstrate the asymmetric effect of the volatility index (*VIX*) on mutual funds' portfolio composition. Positive shocks in the *VIX* increase stock investment in mutual fund portfolios in the long run. The influence of positive shocks (0.056) is larger than that of negative shocks (0.034), displaying that positive shocks have more profound effects than negative shocks. This result suggests that when the *VIX* increases (mutual funds may face more losses), mutual funds invest more money into the stock market and are risk-seeking; additionally, they may react more to potential losses (rising volatility) than potential gains (declining volatility). In the short run, the positive shocks of *VIX* have a significantly negative influence on stock investment. The study indicates that mutual funds tend to decrease their holdings in risk assets in response to an increase in positive *VIX* shocks. This suggests a tendency among mutual funds to be more responsive to a rising *VIX* rather than a declining *VIX*, perceiving these positive shocks as indicative of heightened risks. Consequently, this leads to a reduction in stock investments as a strategy to mitigate potential losses.

Furthermore, the results show that both positive and negative changes in *VIX* impact mutual funds' cash holdings significantly in the long run. Mutual funds will increase the cash amount during positive shocks of *VIX* and decrease it during negative shocks of *VIX*. This implies that mutual funds may increase safe assets for risk management and be more confident in decreasing cash to invest in a stable market environment. However, in the short run, cash holdings are only positively impacted by positive shocks, suggesting that mutual funds may increase their cash assets during a period of increasing *VIX* to avoid risk. In addition, when the market becomes more volatile, mutual fund managers may increase their cash holdings as an insurance to ease the pressure associated with investor redemption without selling underlying assets(Morris et al., 2017).

The above results fit into the prospect theory established by Kahneman and Tversky (1979). Prospect theory is a concept in behavioural finance that posits that people's decision-making is influenced by their perception of potential gains and losses rather than solely by the outcome. The theory proposes that people exhibit a tendency toward loss aversion, meaning that the negative impact of potential losses is greater than that of equivalent gains, and this plays a significant role in their decision-making. The results also conform to the value function proposed in prospect theory. The value function in prospect theory is a graphical representation of the value placed on potential gains and losses. The value function has a key feature: loss aversion, meaning that individuals are more sensitive to possible losses than benefits of equal magnitude. This indicates that the slope of the value function is higher for losses compared to profits.

The above results demonstrate that mutual funds react more to potential losses than potential gains. The negative shocks of *BW* and positive shocks of *VIX* and *GO* represent that investor sentiment is decreasing, and the uncertainty and risk may be detected by investors. Considering the dependent variable *Ps*, the coefficients of $BW^-$ and $VIX^+$ are larger than $BW^+$ and $VIX^-$, respectively. Thus, this suggests that mutual funds' stock investment amount is more sensitive to losses accompanied by $BW^-$ and $VIX^+$. As for cash holdings amount, mutual funds are more sensitive to negative shocks of *BW* in both the long run and short run. In addition, the short-run positive shocks of *GO* and *VIX* have lager coefficients than the negative shocks, which indicates funds are more sensitive to situations with a higher probability of losing (higher volatility). This also indicates that the activities of mutual funds are consistent with prospect theory and reflect the key feature of value function, loss aversion.

However, the long-run relationship between the fluctuation of *CCI*, *VIX*, and cash holdings amounts violate the value function. Mutual funds react more to positive sentiment changes than to negative changes. The negative shocks of the Google search volume of 'economic recession' and 'financial crisis' impact the mutual fund stock investment amount more than positive shocks in the long run, and there is no significant impact in the short run. *GO* may reflect the chance of having a financial crisis; when *GO* stays high, that may suggest the crisis is striking, and vice versa. This implies that in the long run, mutual funds will enter the stock market when the chance of financial crisis is low and be more reactive when *GO* decreases. This is not consistent with value function. However, *GO* has a significantly positive impact in the short run but no significant long-run effect to mutual funds' cash holdings. This suggests mutual funds increase their safe assets during period of increasing *GO* to avoid risks and losses in the short term, and this is explained with prospect theory.

Regarding the BW Index, the stock investments and cash holdings of mutual funds are more sensitive to negative shocks in both the long and short run. $BW^{+}$ has a significantly negative impact on stock investment and positive impact on cash holdings. That implies that during negative shocks, mutual funds may transfer cash to increase their stock investments. This suggests that mutual funds' activities are consistent with the value function; they react more when sentiment decreases (higher risk) and is loss averse. Nevertheless, mutual funds will decrease their cash holdings during negative periods of the BW Index in the short run; this indicates they may prefer a short-term negative shock of sentiment to invest their cash in assets that have become undervalued due to the negative sentiment.

Additionally, the stock investment of mutual funds is more sensitive to the positive shocks of *VIX* in both the long run and short run; expressly, mutual funds tend to adjust their proportion of stock during a period of increasing *VIX* (loss). Especially in the short run, mutual funds will decrease their stock investment during *VIX* positive shocks, whereas the negative shocks have no significant effect. This indicates that mutual funds tend to have higher subjective valuation toward loss and risk. Overall, these results fit into prospect theory.

The asymmetric effects of investor sentiment on mutual fund portfolio composition provide critical insights for fund managers. Understanding that mutual funds react more strongly to negative sentiment changes than to positive ones can significantly influence risk management strategies. During periods of heightened volatility or negative sentiment shocks, managers often increase cash holdings to mitigate potential losses, demonstrating a loss-averse approach. Conversely, in response to positive sentiment shocks, mutual funds tend to allocate more assets to stocks, viewing heightened sentiment as an opportunity to seek higher returns, even at the cost of increased risk. Recognizing these patterns allows managers to develop more balanced asset allocation strategies, aligning the pursuit of gains with effective risk management.

Moreover, acknowledging these asymmetric effects can help mutual fund managers address behavioral biases in decision-making. Awareness of their own potential for loss aversion can encourage managers to adopt a more rational and objective approach, reducing the influence of sentiment-driven fluctuations. This disciplined mindset can enhance investment strategies, curbing the tendency to overreact to market sentiment shifts and promoting more consistent portfolio management practices.

For both individual and institutional investors, understanding the asymmetric effects of sentiment, grounded in principles of prospect theory, can guide better investment decisions and portfolio construction. By recognizing their own loss aversion tendencies, investors can mitigate its impact on their choices, focusing on long-term objectives rather than short-term sentiment fluctuations.

Institutional investors, such as mutual funds, can further leverage sentiment analysis to refine their strategic decision-making processes. By integrating sentiment trends into their investment frameworks, they can anticipate market reactions more effectively and adjust their portfolios proactively. This forward-looking approach can enhance portfolio performance and strengthen risk management, helping investors navigate sentiment-driven market dynamics and periods of heightened volatility with greater confidence.

The identified asymmetric effects of investor sentiment on mutual fund portfolio composition provide valuable insights for mutual fund managers. Understanding that mutual funds are more responsive to negative sentiment changes than positive ones can inform risk management strategies. During periods of rising volatility or negative sentiment shocks, managers might increase cash holdings to mitigate potential losses, reflecting a loss-averse behaviour. Conversely, during periods of positive sentiment shocks, mutual funds tend to increase stock investments. This suggests that managers perceive higher sentiment as an opportunity to seek higher returns, despite the potential for increased risk. Recognizing these behavioural tendencies can help managers devise more effective asset allocation strategies, balancing the pursuit of gains with the need to manage risks.

The asymmetric effects of sentiment can also help mutual fund managers address behavioural biases in investment decision-making. By acknowledging their own potential for loss aversion, managers can strive to make more rational, objective decisions, rather than being swayed excessively by fluctuations in investor sentiment. This understanding can lead to more disciplined investment approaches, reducing the likelihood of overreacting to market sentiment changes.

For both individual and institutional investors, understanding the asymmetric effects of sentiment and the principles of prospect theory can guide investment decisions and portfolio construction. Investors can benefit from recognizing their own loss aversion tendencies and seeking to mitigate its impact on their investment choices. By considering the long-term implications of sentiment fluctuations, investors can better navigate periods of market volatility and sentiment-driven market movements. Institutional investors, such as mutual funds, can utilize these insights to enhance their strategic decision-making processes. By incorporating sentiment analysis into their investment strategies, they can better anticipate market reactions to sentiment changes and adjust their portfolios accordingly. This proactive approach can improve overall portfolio performance and risk management.

As for the control variables, *IR* has a significantly positive impact on mutual fund stock investments in the long run, except for the *GO* and *NIPO* regressions. Furthermore, it has a long-run significantly positive effect on cash holdings in *CCI*, *NIPO*, *RIPO*, and *VIX* regressions. This suggests mutual funds will increase their stock investment and cash holdings in the long run, because mutual funds may regard the increase in the short-term interest rate as positive news or gain the interest income. This is consistent with Ciminelli et al. (2022) conclusion that an increase in the short-term interest rate may indicate a positive economic outlook and growth expectation, which will potentially ease risk-aversion among investors and raise their appetite for risk to increase their stock investment. As for *UEMP*, it has a positive relationship with stock investment and a negative relationship with cash holdings in the long run. This may be due to institutional investors like mutual funds setting long-term investment strategies and entering the stock market with a lower price due to the increasing unemployment rate. Furthermore, the result is consistent with the study byAtanasov (2021), which suggests that a rise in the unemployment rate, relative to its trend, is associated with a rise in higher excess returns. Therefore, investors appear to view increasing unemployment rates as investment opportunities, potentially due to the expectation of higher returns during such periods.

## 4.4 Robustness Check

The investor sentiment indicators might have an asymmetric impact on mutual fund portfolio composition. Therefore, this study applies the Wald test by setting the null hypothesis as $H_0 = \beta^+ = \beta^-$ indicating no asymmetric impact, and the alternative hypothesis is set as $H_a = \beta^+ \neq \beta^-$ implying that an asymmetric impact exists.

**Table 4–4. Wald test**

| | *Ps* | *Decision* | *Pc* | *Decision* |
|---|---|---|---|---|
| ***$CCI_{LR}$*** | 5.26 [0.022] | Yes | 0.96 [0.327] | No |
| ***$CCI_{SR}$*** | 0.11 [0.737] | No | 1.46 [0.227] | No |
| ***$GO_{LR}$*** | 38.01 [0.000] | Yes | 33.95 [0.000] | Yes |
| ***$GO_{SR}$*** | 0.02 [0.892] | No | 7.83 [0.001] | Yes |
| ***$BW_{LR}$*** | 54.93 [0.000] | Yes | 12.10 [0.001] | Yes |
| ***$BW_{SR}$*** | 6.70 [0.010] | Yes | 13.35 [0.000] | Yes |
| ***$NIPO_{LR}$*** | 16.71 [0.000] | Yes | 9.44 [0.002] | Yes |
| ***$NIPO_{SR}$*** | 1.78 [0.182] | No | 0.73 [0.393] | No |
| ***$RIPO_{LR}$*** | 26.25 [0.000] | Yes | 9.20 [0.002] | Yes |
| ***$RIPO_{SR}$*** | 1.39 [0.239] | No | 1.72 [0.190] | No |
| ***$VIX_{LR}$*** | 47.41 [0.000] | Yes | 15.98 [0.000] | Yes |
| ***$VIX_{SR}$*** | 8.06 [0.005] | Yes | 0.41 [0.520] | No |

Notes: Ps and Pc refer to the percentage of the capital amount of mutual funds invested in common stock, and cash, respectively. CCI means the consumer confidence index, GO means google trend searching volume of financial crisis and economic recession, BW indicates Baker and Wurgler index, NIPO means the number of IPOs, RIPO means the first day return of IPOs, VIX means the volatility index. LR and SR represent long run and short run asymmetry. The results show Chin-square and p-value in bracket.

The results of the Wald test are demonstrated in Table 4-3. All investor sentiment measurements have a long-run asymmetric effect on the stock proportion of mutual fund portfolio compositions. Only *BW* and *VIX* have a short-run asymmetry effect on stock investment of mutual funds. Nevertheless, only *CCI* has no significantly asymmetric influence on the cash holdings of mutual funds neither in the long run nor short run. *BW* and *GO* have a significant asymmetric effect in both the long run and short run, whereas the rest only have a long-run asymmetry impact, but the short-run is absent.

## 4.4.1 Dumitrescu-Hurlin Causality Test

The reasons for conducting a causality test are reported in Chapter 3. In the discussions that follow, it is my intention to continuously prompt the reader to consider the estimated causal effects as provisional and not final.

# Table 4–5. Dumitrescu-Hurlin causality test

| | *Ps* | | | *Pc* | |
|---|---|---|---|---|---|
| | Wald-statistics | p-value | | Wald-statistics | p-value |
| *$BW^+$ does not homogeneously cause Ps* | 3.349 | 0.000 | *$BW^+$ does not homogeneously cause Pc* | 4.326 | 0.000 |
| *Ps does not homogeneously cause $BW^+$* | 1.295 | 0.102 | *Pc does not homogeneously cause $BW^+$* | 5.897 | 0.000 |
| *$BW^-$ does not homogeneously cause Ps* | 3.565 | 0.000 | *$BW^-$ does not homogeneously cause Pc* | 5.066 | 0.000 |
| *Ps does not homogeneously cause $BW^-$* | 2.344 | 0.000 | *Pc does not homogeneously cause $BW^-$* | 1.604 | 0.000 |
| *$CCI^+$ does not homogeneously cause Ps* | 3.839 | 0.000 | *$CCI^+$ does not homogeneously cause Pc* | 4.848 | 0.000 |
| *Ps does not homogeneously cause $CCI^+$* | 10.952 | 0.000 | *Pc does not homogeneously cause $CCI^+$* | 0.910 | 0.401 |
| *$CCI^-$ does not homogeneously cause Ps* | 2.990 | 0.000 | *$CCI^-$ does not homogeneously cause Pc* | 4.966 | 0.000 |
| *Ps does not homogeneously cause $CCI^-$* | 1.068 | 0.863 | *Pc does not homogeneously cause $CCI^-$* | 1.748 | 0.000 |
| *$GO^+$ does not homogeneously cause Ps* | 2.717 | 0.000 | *$GO^+$ does not homogeneously cause Pc* | 6.760 | 0.000 |
| *Ps does not homogeneously cause $GO^+$* | 4.622 | 0.000 | *Pc does not homogeneously cause $GO^+$* | 6.760 | 0.000 |
| *$GO^-$ does not homogeneously cause Ps* | 3.121 | 0.000 | *$GO^-$ does not homogeneously cause Pc* | 6.507 | 0.000 |
| *Ps does not homogeneously cause $GO^-$* | 3.095 | 0.000 | *Pc does not homogeneously cause $GO^-$* | 2.101 | 0.000 |
| *$NIPO^+$ does not homogeneously cause Ps* | 2.988 | 0.000 | *$NIPO^+$ does not homogeneously cause Pc* | 3.452 | 0.000 |
| *Ps does not homogeneously cause $NIPO^+$* | 0.752 | 0.064 | *Pc does not homogeneously cause $NIPO^+$* | 1.394 | 0.023 |
| *$NIPO^-$ does not homogeneously cause Ps* | 4.311 | 0.000 | *$NIPO^-$ does not homogeneously cause Pc* | 4.045 | 0.000 |
| *Ps does not homogeneously cause $NIPO^-$* | 1.039 | 0.992 | *Pc does not homogeneously cause $NIPO^-$* | 0.861 | 0.247 |
| *$RIPO^+$ does not homogeneously cause Ps* | 3.770 | 0.000 | *$RIPO^+$ does not homogeneously cause Pc* | 4.345 | 0.000 |
| *Ps does not homogeneously cause $RIPO^+$* | 0.921 | 0.000 | *Pc does not homogeneously cause $RIPO^+$* | 0.888 | 0.326 |
| *$RIPO^-$ does not homogeneously cause Ps* | 4.088 | 0.000 | *$RIPO^-$ does not homogeneously cause Pc* | 4.368 | 0.000 |
| *Ps does not homogeneously cause $RIPO^-$* | 0.808 | 0.135 | *Pc does not homogeneously cause $RIPO^-$* | 1.074 | 0.831 |
| *$VIX^+$ does not homogeneously cause Ps* | 2.988 | 0.000 | *$VIX^+$ does not homogeneously cause Pc* | 4.568 | 0.000 |
| *Ps does not homogeneously cause $VIX^+$* | 1.189 | 0.340 | *Pc does not homogeneously cause $VIX^+$* | 1.077 | 0.815 |
| *$VIX^-$ does not homogeneously cause Ps* | 3.383 | 0.000 | *$VIX^-$ does not homogeneously cause Pc* | 5.400 | 0.000 |
| *Ps does not homogeneously cause $VIX^-$* | 1.356 | 0.043 | *Pc does not homogeneously cause $VIX^-$* | 1.113 | 0.644 |

* Ps and Pc refer to the percentage of capital amount of mutual funds invested in common stock, and cash, respectively. CCI means the consumer confidence index, GO means google trend searching volume of financial crisis and economic recession, BW indicates Baker and Wurgler index, NIPO means the number of IPOs, RIPO means the first day return of IPOs, VIX means the volatility index. + indicates positive change, - indicates negative change.

As Table 4-4 illustrates, Dumitrescu-Hurlin causality tests indicate some unidirectional causality from decomposed investor sentiment indicators to both categories of mutual funds portfolio composition. The tests show unidirectional causality from $BW^{+}$, $CCI^{-}$, $NIPO^{-}$, $RIPO^{-}$, and $VIX^{+}$ to $Ps$; and from $CCI^{+}$, $NIPO^{-}$, $RIPO^{+}$, $RIPO^{-}$, $VIX^{+}$, and $VIX^{-}$ to $Pc$. These results indicate that the mutual fund portfolio composition is impacted by the fluctuation of multiple investor sentiment proxies.

### 4.4.2 Sub-Sample Regressions

This chapter conducts the same sub-sample period regressions as Chapter 3, using the relatively stable period of 2010 Q1 to 2019 Q4 to detect whether the financial crisis is crucial to the results.

## Table 4–6. Sub-sample nonlinear panel ARDL results of Ps

| *Variables* | *Ps* | | | | | |
|---|---|---|---|---|---|---|
| | **(a)** | **(b)** | **(c)** | **(d)** | | |
| | ***Survey* base** | ***Internet base*** | ***Composite*** | ***Market base*** | | |
| | | | ***Long run equation*** | | | |
| ***CCI*** $^{+}$ | 0.033 | | | | | |
| | (0.019) | | | | | |
| ***CCI*** $^{-}$ | 0.065* | | | | | |
| | (0.021) | | | | | |
| ***GO*** $^{+}$ | | -0.273* | | | | |
| | | (0.048) | | | | |
| ***GO*** $^{-}$ | | -0.018 | | | | |
| | | (0.046) | | | | |
| ***BW*** $^{+}$ | | | -0.480 | | | |
| | | | (0.955) | | | |
| ***BW*** $^{-}$ | | | 1.881* | | | |
| | | | (0.862) | | | |
| ***NIPO*** $^{+}$ | | | | 0.028* | | |
| | | | | (0.008) | | |
| ***NIPO*** $^{-}$ | | | | 0.028* | | |
| | | | | (0.005) | | |
| ***RIPO*** $^{+}$ | | | | | 0.003 | |
| | | | | | (0.001) | |
| ***RIPO*** $^{-}$ | | | | | 0.020 | |
| | | | | | (0.002) | |
| ***VIX*** $^{+}$ | | | | | | -0.078* |
| | | | | | | (0.025) |
| ***VIX*** $^{-}$ | | | | | | 0.047 |
| | | | | | | (0.025) |
| ***IR*** | -0.369 | 0.350* | 0.593* | 0.269 | 0.354* | 0.406* |
| | (0.211) | (0.146) | (0.297) | (0.219) | (0.018) | (0.149) |
| ***UEMP*** | -0.168 | -1.315* | -0.602* | -0.134 | -0.463* | -1.170* |
| | (0.389) | (0.276) | (0.232) | (0.266) | (0.017) | (0.175) |
| ***GPR*** | -0.050 | -0.067 | 0.436 | 0.807* | 0.436 | -0.378 |
| | (0.288) | (0.265) | (0.267) | (0.262) | (0.261) | (0.254) |
| | | | ***Short run equation*** | | | |
| ***\|COINTEQ*** | -0.347* | -0.368* | -0.350* | -0.342* | -0.353* | -0.359* |
| | (0.027) | (0.028) | (0.028) | (0.028) | (0.028) | (0.029) |
| ***D(CCI*** $^{+}$***)*** | -0.019 | | | | | |
| | (0.021) | | | | | |
| ***D(CCI*** $^{-}$***)*** | -0.008 | | | | | |
| | (0.021) | | | | | |
| ***D(GO*** $^{+}$***)*** | | 0.032 | | | | |
| | | (0.062) | | | | |
| ***D(GO*** $^{-}$***)*** | | 0.046 | | | | |
| | | (0.036) | | | | |
| ***D(BW*** $^{+}$***)*** | | | 0.185 | | | |
| | | | (0.884) | | | |
| ***D(BW*** $^{-}$***)*** | | | -0.356 | | | |
| | | | (1.180) | | | |
| ***D(NIPO*** $^{+}$***)*** | | | | -0.002 | | |
| | | | | (0.006) | | |
| ***D(NIPO*** $^{-}$***)*** | | | | -0.001 | | |
| | | | | (0.006) | | |
| ***D(RIPO*** $^{+}$***)*** | | | | | -0.037* | |
| | | | | | (0.013) | |
| ***D(RIPO*** $^{-}$***)*** | | | | | 0.028* | |
| | | | | | (0.028) | |
| ***D(VIX*** $^{+}$***)*** | | | | | | -0.016 |
| | | | | | | (0.019) |
| ***D(VIX*** $^{-}$***)*** | | | | | | 0.044 |
| | | | | | | (0.030) |
| ***D(IR)*** | -0.428 | -0.510 | -0.476 | -0.170 | -0.462 | -0.713 |
| | (0.443) | (0.481) | (0.434) | (0.476) | (0.466) | (0.468) |
| ***D(UEMP)*** | 0.218 | 0.587 | 0.530 | 0.588 | 0.357 | 0.579 |
| | (0.367) | (0.339) | (0.364) | (0.351) | (0.348) | (0.365) |
| ***D(GPR)*** | 0.069 | 0.039 | -0.074 | -0.123 | -0.153 | 0.112 |
| | (0.112) | (0.036) | (0.118) | (0.122) | (0.115) | (0.116) |
| ***C*** | 30.638* | 36.791* | 31.731* | 29.093* | 31.667* | 29.762* |
| | (2.623) | (3.063) | (2.770) | (2.578) | (2.748) | (2.631) |

Notes: * denotes statistically significant at 5% level. The results demonstrate coefficient and standard error with parentheses. + indicates positive change, - indicates negative change, Ps and Pc refer to the percentage of capital amount of mutual funds invested in common stock, and cash, respectively. CCI means the consumer confidence index, GO means google trend searching volume of financial crisis and economic recession, BW indicates Baker and Wurgler index, NIPO means the number of IPOs, RIPO means the first day return of IPOs, VIX means the volatility index, IR means short-term interest rate, UEMP indicates the unemployment rate, and GPR refers to the geopolitical risk index.

## Table 4–7. Sub-sample nonlinear panel ARDL results of Pc

| *Variables* | *Pc* (a) *Survey* base | (b) *Internet base* | (c) *Composite* | (d) *Market base* | | |
|---|---|---|---|---|---|---|
| | | | *Long run equation* | | | |
| $CCI^{+}$ | 0.005* (0.002) | | | | | |
| $CCI^{-}$ | -0.0002 (0.002) | | | | | |
| $GO^{+}$ | | -0.0004 (0.005) | | | | |
| $GO^{-}$ | | -0.012* (0.005) | | | | |
| $BW^{+}$ | | | 0.121 (0.093) | | | |
| $BW^{-}$ | | | 0.130 (0.085) | | | |
| $NIPO^{+}$ | | | | 0.002* (0.001) | | |
| $NIPO^{-}$ | | | | 0.0002 (0.0004) | | |
| $RIPO^{+}$ | | | | | 0.002 (0.001) | |
| $RIPO^{-}$ | | | | | 0.001 (0.002) | |
| $VIX^{+}$ | | | | | | 0.004 (0.003) |
| $VIX^{-}$ | | | | | | 0.002 (0.003) |
| *IR* | -0.002 (0.018) | 0.022 (0.014) | 0.027 (0.029) | -0.022 (0.205) | 0.006 (0.015) | 0.014 (0.016) |
| *UEMP* | 0.102* (0.034) | 0.079* (0.027) | 0.015 (0.022) | 0.061* (0.025) | 0.032* (0.015) | -0.027* (0.019) |
| *GPR* | 0.065* (0.025) | 0.082* (0.026) | 0.057* (0.263) | 0.054* (0.025) | 0.063* (0.023) | 0.058* (0.029) |
| | | | *Short run equation* | | | |
| *\|COINTEQ* | -0.631* (0.029) | -0.651* (0.030) | -0.631* (0.029) | -0.622* (0.029) | -0.628* (0.029) | -0.647* (0.028) |
| $D(CCI^{+})$ | -0.033* (0.015) | | | | | |
| $D(CCI^{-})$ | -0.027 (0.014) | | | | | |
| $D(GO^{+})$ | | 0.044 (0.045) | | | | |
| $D(GO^{-})$ | | -0.046 (0.031) | | | | |
| $D(BW^{+})$ | | | 0.274 (0.800) | | | |
| $D(BW^{-})$ | | | -2.606* (1.187) | | | |
| $D(NIPO^{+})$ | | | | 0.004 (0.005) | | |
| $D(NIPO^{-})$ | | | | -0.001 (0.003) | | |
| $D(RIPO^{+})$ | | | | | -0.001 (0.011) | |
| $D(RIPO^{-})$ | | | | | 0.007 (0.008) | |
| $D(VIX^{+})$ | | | | | | 0.041* (0.025) |
| $D(VIX^{-})$ | | | | | | -0.039 (0.025) |
| *D(IR)* | -1.221* (0.461) | -1.214* (0.470) | -1.510* (0.484) | -1.199* (0.448) | -1.224 (0.410) | -1.500* (0.484) |
| *D(UEMP)* | -0.635 (0.355) | -0.212 (0.333) | -0.092 (0.331) | -0.126 (0.314) | -0.158 (0.314) | -0.303 (0.309) |
| *D(GPR)* | 0.233 (0.119) | 0.150 (0.098) | 0.065 (0.104) | 0.144 (0.101) | 0.113 (0.099) | 0.138 (0.105) |
| *C* | 0.148* (0.140) | 0.085* (0.165) | 0.595* (0.143) | 0.375* (0.140) | 0.594* (0.131) | 0.489* (0.132) |

Notes: * denotes statistically significant at 5% level. The results demonstrate coefficient and standard error with parentheses. + indicates positive change, - indicates negative change, Ps and Pc refer to the percentage of capital amount of mutual funds invested in common stock, and cash, respectively. CCI means the consumer confidence index, GO means google trend searching volume of financial crisis and economic recession, BW indicates Baker and Wurgler index, NIPO means the number of IPOs, RIPO means the first day return of IPOs, VIX means the volatility index, IR means short-term interest rate, UEMP indicates the unemployment rate, and GPR refers to the geopolitical risk index.

According to Table 4–5 and Table 4–6, excluding *NIPO* and *RIPO*, the investor sentiment indices may have an asymmetrical relationship with mutual funds' amount of stock investment in the long run. The dependent variable *Ps* has a significant relationship with the negative shocks of *CCI* and *BW* and the positive shocks of *GO* and *VIX*. Each implies that the market may suffer losses. Therefore, mutual funds are more sensitive and have higher subjective valuation toward losses than gains. Thus, the results further confirm that the activities of mutual funds are consistent with the value function of prospect theory. However, investor sentiment only has a limited effect on the cash holdings of mutual funds during a stable time and cannot be explained with prospect theory. This may be explained by mutual funds not adjusting their safe assets frequently when the market condition is stable. In addition, the results show that mutual funds react more to negative shocks of investor sentiment during stable times. Although, in some cases, they react more to positive shocks of sentiment during the period of crises, which is not aligned with prospect theory. The reason for this phenomenon may be that investors are more eager to chase positive information when crisis strike, and the situation could not be worse; which leading them react to any good news. During a stable time, mutual funds show the feature of loss aversion and react more to negative information or negative shocks of sentiment. Overall, the results suggest that the prospect theory applies to intuitional investors' activities like mutual funds.

# 4.5 Conclusion

This chapter discussed the nonlinear relationship between different sentiment indicators and the portfolio composition of mutual funds. The results demonstrate that both positive and negative changes will impact portfolio composition overall. Furthermore, the stock proportion of mutual funds has an asymmetric nexus with investor sentiment measurements. However, the fluctuation of investor sentiment has a limited asymmetric effect on the cash holdings of mutual funds, except for the BW Index, which has a larger coefficient.

In conclusion, positive and negative changes in investor sentiment appear to have asymmetric impacts on the allocation of mutual fund assets between common stocks and cash. The prospect theory can be incorporated into the analysis and explain the asymmetry effect from investor sentiment indices, such as the BW Index and volatility index, on mutual fund portfolio composition. In addition, the results show that mutual funds are more consistent with the value function or loss aversion during a stable time than a crisis period. This study provides a more nuanced understanding of the relationship between investor sentiment and mutual fund portfolio allocation, as it considers how investors perceive gains and losses and make decisions based on subjective values. It also highlights the importance of considering behavioural biases in investment decision-making and portfolio management. These findings may have important implications for investing decisions and portfolio management strategies.

# Chapter 5. A New Composite Investor Sentiment Index: Does It Affect Mutual Funds' Portfolio Composition?

## 5.1 Introduction

In the realm of financial economics, the construct of investor sentiment is recognized as a pivotal determinant in the dynamics of financial markets. Investor sentiment, broadly defined as the attitudes and expectations of investor toward market conditions, exerts a substantial influence on asset pricing and market behaviour. This sentiment is inherently complex, shaped by a confluence of diverse factors, such as macroeconomic indicators, market news, and sociopolitical events. Consequently, an accurate assessment of investor sentiment necessitates a multifactorial approach, transcending the limitations of single-indicator analyses.

It could be true that individual sentiment measures would have its own set of limitations, biases, and area of coverage., since they often tap into only a small slice of the overall sentiment landscape. For instance, survey-based index can obtain individual investors' opinion directly but may omit the potential market fluctuation, while market-based indices reflect trading behaviour but not subjective attitudes. As for internet-based index, it can only capture the concerning point of individuals during one period, which may contain noise and not reveal the real condition of the market,

but it sure can be regarded as a reflection of investors' thoughts. True sentiment is a theoretical construct that reflects the overall psychological state of the market or investors, which we cannot directly observe.

To address these limitations, composite indices combine multiple sentiment measures to approximate true sentiment more effectively. By integrating diverse sources, they reduce noise and idiosyncratic biases, creating a fuller representation of sentiment. That said, composite indices are still approximations—they depend on the validity of their components and can lose granularity or introduce biases through aggregation. It is hard to observe true sentiment directly, composite indices offer a practical and theoretically sound approach to approximate it, using the best available data from multiple proxies.

In this context, a principal component analysis (PCA) stands out as an efficacious statistical methodology for the construction of a composite sentiment index. The PCA, renowned for its capacity to reduce dimensionality and highlight significant patterns within a dataset, is particularly suited for concentrating various sentiment indicators into a coherent index. This technique not only simplifies the complexity inherent in multiple data sources but also augments the interpretability and analytical utility of the resultant sentiment index.

The seminal work by Baker and Wurgler (2006; 2007) exemplifies the application of PCA. Their methodology adeptly integrates six distinct sentiment proxies – encompassing metrics such as the closed-end fund discount, NYSE share turnover, the volume of IPOs, first-day returns of IPOs, the equity-to-debt issue ratio, and the dividend premium – into an aggregate sentiment index. This index accounted for a substantial proportion of the variance among the individual proxies, thereby demonstrating the effectiveness of PCA in synthesizing a composite sentiment measure.

Notwithstanding, the construction of a composite sentiment index via PCA is replete with analytical challenges. Each sentiment proxy is likely to contain both a sentiment component and non-sentiment-related components. A principal component analysis is used to identify the common component. In addition, understanding the temporal dynamics and lead-lag correlations among these variables is imperative for accurately capturing and interpreting shifts in investor sentiment.

In this chapter, the objective is to extend and refine the methodology pioneered by Baker and Wurgler (2006; 2007), by developing an advanced composite sentiment index utilizing a two-step PCA approach. This endeavour is a critical step toward a deeper comprehension of market mechanisms and the integral role of investor sentiment. By devising a more encompassing and precise sentiment index, this study aspires to enhance the understanding of market behaviours, thereby contributing to more informed investment decision-making and economic policy formulation.

Compared to earlier studies, this chapter introduces a single PCA-based composite index with the integration of contemporary and diverse indicators, which include survey-based, market-based and Internet-based sentiment indices, and the proposed index demonstrates an expanded analytical scope by capturing a more comprehensive array of sentiment drivers. While previous research predominantly focused on sentiment's influence on stock returns and market volatility(Baker and Wurgler, 2006; Huang et al., 2015). The current study shifts the focus to how sentiment influences institutional investment behaviour. By combining a broader range and different kinds of sentiment indicators and applying Baker and Wurgler (2006)'s PCA methodology, the proposed index provides a more nuanced and predictive measure of how shifts in investor sentiment translate into portfolio reallocation decisions. Compared to other research, the main different of this thesis is attempt to include more different based sentiment indices inside rather than only consider market-based proxies. Consequently, this study contributes to the literature by offering a more detailed and multifaceted tool for understanding the relationship between investor sentiment and mutual fund behaviour, paving the way for more informed asset management strategies. Furthermore, by directly linking this sentiment index to mutual fund portfolio composition—specifically, the allocation between stocks and cash holdings—this study offers new insights into the behavioural dynamics of institutional investors, an area that has been less explored by Liao et al. (2011), who construct one PCA index to detect the herding activity of mutual funds. In summary, this study employs an enhanced methodology to demonstrate that the new (aligned investor sentiment) index possesses higher correlation with mutual fund portfolio composition than the original metric developed by Baker and Wurgler (2006; 2007). This advancement is achieved by integrating a broader array of sentiment indicators, thereby encapsulating a more comprehensive spectrum of market information. The research conclusively establishes that this novel index effectively assimilates and utilizes the information inherent in the five selected proxies, and they can be classified into the survey-based, like CCI, market-based like NIPO, RIPO, VIX and BW index, and internet-based like GO, thereby offering a more robust tool for reflecting the composition of mutual fund portfolios.

This chapter carries on in four sections. The next section outlines the data sample and methodology. Section 5.3 demonstrate the empirical results. Section 5.4 illustrate the robustness check. And the last section provides conclusion.

## 5.2 Data and Methods

This chapter utilizes the dependent variables established in the preceding chapter, building upon the empirical findings that highlight the substantial influence of both survey-based and internet-based proxies on mutual fund portfolio composition. This observation underscores a critical limitation in the BW Index, which is exclusively composed of market-based proxies. The BW Index may potentially overlook the rich insights offered by non-market data sources. My new index addresses this gap by integrating a more diverse set of proxies, including those derived from surveys and internet sources, thereby offering a more holistic view of investor sentiment. This comprehensive approach is designed to capture a broader spectrum of market influences, providing a more accurate tool for analysing mutual fund portfolio dynamics. Therefore, some proxies are considered in the construction of the investor sentiment index. The consumer sentiment index (*CSI*) and the consumer confidence index (*CCI*) are proxies for survey-based investor sentiment measurement. In addition, the number of IPOs issued (*NIPO*) and the volatility index (*VIX*) serve as market-based investor sentiment metrics. Finally, the Google search volume of the keywords ‘financial crisis’ and ‘economic recession’ is the internet-based sentiment index. I use quarterly data for the period from 2007 Q3 to 2021 Q2. The dataset is the same as in the previous chapters.

### 5.2.1 Creating the Sentiment Index

This research first applies a PCA to construct an investor sentiment indicator. PCA is a technique for dimension reduction. It is widely used in many empirical studies. Principal Component Analysis (PCA) is a statistical method that applies an orthogonal transformation to transform a dataset containing possibly correlated variables into a new set of linearly uncorrelated variables known as principal components. This transformation is structured so that the first principal component captures the maximum variance possible (meaning it explains the largest part of the variability in the dataset), and each subsequent component, while being orthogonal to its predecessors, captures the next highest amount of variance possible.

The goal of PCA is to identify the directions (or principal components) where the data varies the most, as these directions often contain the most 'information' about the differences between observations. In a PCA, the first principal component is chosen in such a way that it accounts for the largest possible variance in the dataset. Expressly, the first principal component is the linear combination of the original variables (i.e., features) that captures the most variance in the data. The second principal component is calculated in the same way, with the condition that it is orthogonal (i.e., uncorrelated) to the first principal component and captures the second most variance, and so on.

By focusing on the directions of maximum variance, a PCA allows us to extract the essential features of the dataset, reducing its dimensionality without losing important information. As for eigenvalues, each principal component corresponds to an eigenvalue of the covariance matrix of the original data. The eigenvalue associated with each principal component describes the amount of variance in the data that is accounted for by that component. Expressly, larger eigenvalues correspond to larger amounts of variance. Therefore, the principal components are often ordered by their corresponding eigenvalues.

Overall, the PCA method is employed due to its advantages. The first is its ability to capture the maximum variance in the data with fewer components. PCA achieves this by transforming the original correlated variables into a new set of uncorrelated components (principal components), ordered by the amount of variance each explains in the data set. The first few components typically capture most of the variance, allowing for a reduction in dimensionality without significant loss of information(Jolliffe and Cadima, 2016).Therefore, I can extract more concentrate information from different sentiment proxies. Secondly, PCA is particularly effective in reducing the complexity of such datasets. It also helps eliminate noise and redundancy in the data. By focusing on the principal components, it is possible to ignore components that contribute little to the variation in the data, which often represent noise. Especially, the noises hiding behind different measure of investor sentiment. By concentrating the information contained in the original variables into a smaller number of principal components, PCA simplifies the structure of the data, making it easier to interpret and analyse. It is easier to understand investor sentiment as a whole index rather look at multiple different proxies. And this is especially valuable when dealing with high-dimensional data, where overfitting and multicollinearity are common issues. Thirdly, PCA is adept at identifying underlying factors or components that drive the variance in the data. These factors often correspond to latent variables that are not directly observable but influence multiple observed variables. In the context of investor sentiment, PCA can help identify broad sentiment trends that affect various market indicators, providing a more coherent and unified measure of sentiment (Barberis et al., 1998).

However, PCA also has its limitations. While PCA is effective at reducing dimensionality, it can also lead to the loss of information. This occurs because PCA focuses on maximizing variance, which may not always align with the most meaningful or relevant aspects of the data. For instance, smaller components that capture less variance might still contain valuable information about specific aspects of sentiment, which could be lost when these components are discarded(Ringnér, 2008).. In addition, the decision of how many principal components to retain is somewhat subjective, and different criteria (such as the Kaiser criterion or the scree plot) can lead to different choices. This subjectivity can introduce bias into the analysis, affecting the final composite index. Moreover, retaining too many components may reintroduce complexity, while retaining too few might omit important information(Zwick and Velicer, 1986).

There are alternatives methods of constructing composite index, like factor analysis (FA), data envelopment analysis (DEA) and machine learning (ML) techniques. While both FA and PCA are dimension reduction techniques, FA is more suitable when the goal is to uncover latent constructs that explain the relationships between variables. PCA, on the other hand, is preferable when the primary objective is to maximize variance explained with fewer components (extracting composite index) and when less emphasis is placed on the theoretical underpinning of the components. Then, DEA offers a flexible, non-parametric approach that does not require distributional assumptions, making it useful in diverse datasets. However, its sensitivity to outliers and lack of statistical inference makes it less robust in certain contexts compared to PCA, which provides clear criteria for component retention and interpretation. At last, ML techniques are powerful for predictive modelling, which are beyond PCA's scope. However, ML models are more complex, harder to interpret, and prone to overfitting, whereas PCA offers a simpler, more transparent method that is easier to implement and understand. This makes PCA a preferred choice for many researchers, where interpretability and transparency are highly valued(Abdi and Williams, 2010).

Therefore, regardless the drawback of PCA techniques, this research utilizes a two-stage PCA methodology, following the research of Baker and Wurgler (2006), to construct an investor sentiment index. Investor sentiment is notoriously difficult to quantify due to its subjective nature and the diverse sources from which it can be inferred by various psychological and economic factors. These sources include survey-based measures (e.g., consumer confidence index and consumer sentiment index), market-based indicators (e.g., the number of IPOs and the CBOE volatility index), and internet-based metrics (e.g., Google trend search volume for financial crises and economic recessions). Survey-based sentiment measures, such as the Consumer Confidence Index (CCI) and Consumer Sentiment Index (CSI), offer direct insights into public perceptions and expectations about the economy. These indices are widely used because they capture the mood of consumers, which can significantly influence market behaviour. However, they are subject to limitations, including response biases, the framing of survey questions, and the time lag between data collection and publication(Ludvigson, 2004). Additionally, survey-based measures may not fully capture real-time shifts in sentiment or the more nuanced aspects of investor psychology.

Market-based sentiment measures, such as the number of IPOs and the CBOE Volatility Index (VIX), provide indirect indicators of market sentiment through observable financial market activities. The VIX, often referred to as the "fear gauge," reflects market expectations of near-term volatility and is frequently used as a proxy for investor anxiety or optimism. The number of IPOs can signal investor confidence in the equity markets, with a higher number indicating bullish sentiment. While market-based measures are valuable for their objectivity and real-time data availability, they are influenced by factors beyond sentiment, such as liquidity conditions, regulatory changes, and macroeconomic trends, potentially confounding the interpretation of pure sentiment(Baker and Wurgler, 2007).

Internet-based measures, like Google search volume for terms related to financial crises and economic recessions, offer a novel approach to gauging sentiment by capturing public interest and concern in real-time. These metrics can reflect the collective anxiety or optimism of a broad population segment, extending beyond traditional financial market participants. However, Internet-based indices are prone to noise and may be influenced by non-financial events or media coverage, which can distort the measurement of investor sentiment(Bollen et al., 2011). Moreover, the interpretation of search volume data requires careful contextualization to avoid overestimating the significance of short-term spikes in search activity.

The inclusion of these diverse sentiment proxies in the composite index aims to create a more comprehensive measure of investor sentiment, capturing its various dimensions and manifestations. However, each proxy carries potential limitations and biases that must be acknowledged. For instance, the VIX may be driven by technical market factors unrelated to sentiment, while survey-based indices may lag in reflecting sudden sentiment shifts. Similarly, Google search volume might capture heightened public concern, but it could also be influenced by sensationalist news coverage rather than genuine economic fear. By combining these different measures through PCA, the composite index seeks to mitigate the individual limitations of each proxy and provide a more balanced and robust reflection of overall investor sentiment, but researchers must remain cautious about the interpretative challenges posed by each data source.

This section will elaborate the process of PCA employed in this chapter. Initially, a first-stage index is created, incorporating both current and lagged values of chosen sentiment proxies, resulting in several pairs. Subsequently, the coefficients of each pair are compared, and the proxy with the lower coefficient in each pair is excluded. Finally, the remaining variables, each possessing a higher coefficient, are employed in the second stage PCA analysis. This approach considers that sentiment takes longer to affect the market by including both current and lagged value of all sentiment indicators. Further, it captures the temporal dynamics and relationships between variables over time. Therefore, this procedure is especially useful when investigating investor sentiment.

In addition, it provides information about the impact of past sentiment on current or future situations. According to Baker and Wurgler (2006), the first stage index is the first principal component of all the leads and first lags of the indicators. In their analysis, however, the cumulative variance of the first principal component is only 49%, resulting in the loss of some information from original data.

Consequently, subsequent research by Ding et al. (2017) and Zhengke et al. (2020) enhance this method by establishing criteria to increase the number of components used to produce the first-stage investor sentiment indicator. These studies establish a statistical benchmark of 85% of the cumulative variance rate. Therefore, I selected the first three principal components with a cumulative explained variance rate of 86.967%.

**Table 5–1. Results of first-step principal component analysis**

| | Eigenvalue | Percent variance explained (%) | Cumulative percent explained (%) |
|---|---|---|---|
| F1 | 5.543 | 55.429 | 55.429 |
| F2 | 1.774 | 17.737 | 73.166 |
| F3 | 1.380 | 13.801 | 86.967 |
| F4 | 0.490 | 4.896 | 91.862 |
| F5 | 0.294 | 2.945 | 94.807 |
| F6 | 0.216 | 2.160 | 96.967 |
| F7 | 0.126 | 1.264 | 98.232 |
| F8 | 0.101 | 1.006 | 99.237 |
| F9 | 0.060 | 0.602 | 98.840 |
| F10 | 0.016 | 0.160 | 100.00 |
| Kaiser-Meyer-Olkin measure of sampling adequacy | 0.759 | | |

The results of the first-stage principal component analysis are demonstrated in Table 5–1. The Kaiser-Meyer-Olkin value of sampling adequacy is 0.759, which is above 0.5, indicating that these factors fit into the principal components analysis. The first-stage investor sentiment indicator was then calculated by weighting the three selected principal components by their eigenvalues. Next, I calculated the correlation between first-stage index and each of the 10 variables, the current and lagged value of five sentiment indicators. This procedure captures the common component in the selected sentiment indices that is driving the variation in the dataset, which reflects the common trends and patterns.

Hence, five variables with higher correlation coefficient are selected, and consequently, the rest of variables with lower correlation coefficients were excluded from next stage analysis. The procedures are demonstrated in Figure 5–1.

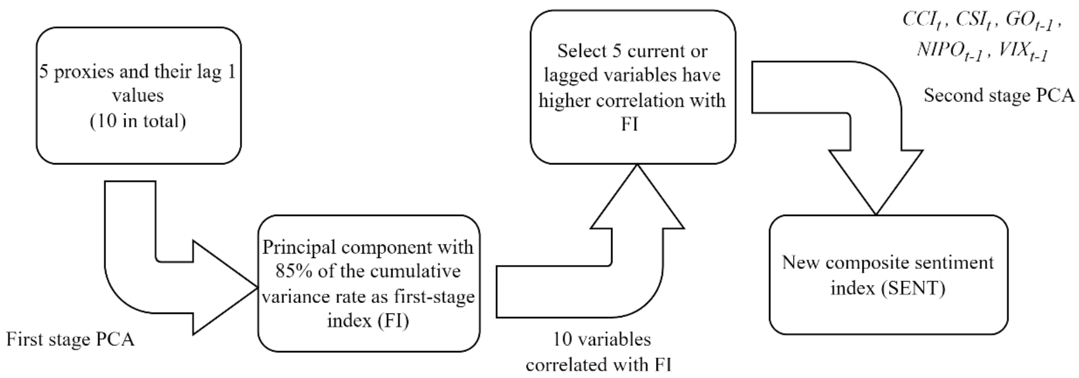



**Figure 5–1. Constructing the investor sentiment index**

**Table 5–2.The comparison of coefficients**

| $CCI_t$ | $CSI_t$ | $GO_t$ | $NIPO_t$ | $VIX_t$ |
|---|---|---|---|---|
| **0.843** | **0.760** | -0.121 | 0.268 | -0.062 |
| $CCI_{t-1}$ | $CSI_{t-1}$ | $GO_{t-1}$ | $NIPO_{t-1}$ | $VIX_{t-1}$ |
| 0.824 | 0.752 | **-0.265** | **0.281** | **-0.304** |

The result of correlation comparison is presented in

Table 5–2. The selected variables are current value of $CCI_t$ and $CSI_t$, and the lagged value of $GO_{t-1}$, $NIPO_{t-1}$ and $VIX_{t-1}$. They will be chosen for the second-step principal components analysis.

**Table 5–3. Results of second step PCA**

| | Eigenvalue | Percent variance explained (%) | Cumulative percent explained (%) |
|---|---|---|---|
| F1 | 2.935 | 58.708 | 58.708 |
| F2 | 0.982 | 19.641 | 78.349 |
| F3 | 0.722 | 14.434 | 92.783 |
| F4 | 0.296 | 5.911 | 98.694 |
| F5 | 0.065 | 1.306 | 100.000 |
| Kaiser-Meyer-Olkin measure of sampling adequacy | | 0.603 | |

Table 5–3 demonstrates that the Kaiser-Meyer-Olkin value of sampling adequacy is 0.603, which is also greater than 0.5, indicating that these factors still fit into the principal components analysis. Then, the eigenvalues and the percentage of variance of the five variables are obtained through dividing the respective eigenvalue

by the sum of all eigenvalues. Next, the weights of each component can be determined. As shown in Table 5–3, the outcomes of the second stage PCA reveal that only the first principal component (*SENT*) is retained. It explains 58.7% variance, indicating the first principal component retains more information than the BW Index.

As illustrated in Table 5–3, among five initial eigenvalues, the value of the first component is greater than one, meaning that its explanatory strength is better than the original variables. Therefore, it is extracted to be the principal components. In addition, the percentage of its contribution to explain the dependent variable is approximately 58.708% cumulatively.

The value of 'percent variance explained' shows that F1 explains 58.708% of the information the original variables contain; F2 explains 19.641%; F3 explains 14.434%; F4 explains 5.911%, and F5 explains 1.306%. The value of 'cumulative percent explained' indicates that the five factors together explain 100% of the information the data contains.

**Table 5–4. Component score coefficient matrix**

| Variables | $CCI_t$ | $CSI_t$ | $GO_{t-1}$ | $NIPO_{t-1}$ | $VIX_{t-1}$ |
|---|---|---|---|---|---|
| Coefficient | 0.301 | 0.305 | -0.264 | 0.108 | -0.275 |

The coefficients of each selected proxy are presented in Table 5–4. Thereby, Equation (5-1) shows the computation of the new sentiment index (*SENT*), and the coefficient of each proxy is obtained through the component coefficient matrix of PCA analysis.

$$SENT_t = 0.301CCI_t + 0.305CSI_t - 0.264GO_{t-1} + 0.108NIPO_{t-1} - 0.275VIX_{t-1} \tag{5-1}$$

### 5.2.2 Model Specification

After the construction of the PCA sentiment index, this research uses the panel ARDL model to examine the power of the SENT index to reflect the mutual fund portfolio composition and their prospect toward risk. The model presented is as follows:

$$\begin{aligned} Pa_{j,t} = {} & \alpha_0 + \beta_{1,j} Pa_{j,t-1} + \beta_2 SENT_{t-1} + \beta_3 IR_{t-1} + \beta_4 UEMP_{t-1} + \beta_5 GPR_{t-1} + \\ & \sum_{i=1}^{p-1} \lambda_{1,j}\, \Delta Pa_{j,t-i} + \sum_{i=0}^{q-1} \lambda_2 \Delta SENT_{t-i} + \sum_{i=0}^{q-1} \lambda_3\, \Delta IR_{t-i} + \sum_{i=0}^{q-1} \lambda_4\, \Delta UEMP_{t-i} + \\ & \sum_{i=0}^{q-1} \lambda_5\, \Delta GPR_{t-i} + \mu_t + \varepsilon_{j,t} \end{aligned} \tag{5-2}$$

According to Pesaran et al. (2001), Equation (5-2) could be reparametrized to an error correction model as follows:

$$\begin{aligned} \Delta Pa_{j,t} = {} & \phi_j ECT_{j,t-1} + \sum_{i=1}^{p-1} \lambda_{1,j}\, \Delta Pa_{j,t-i} + \sum_{i=0}^{q-1} \lambda_2 \Delta SENT_{t-i} + \sum_{i=0}^{q-1} \lambda_3\, \Delta IR_{t-i} + \\ & \sum_{i=0}^{q-1} \lambda_4\, \Delta UEMP_{t-i} + \sum_{i=0}^{q-1} \lambda_5\, \Delta GPR_{t-i} + \mu_t + \varepsilon_{j,t} \end{aligned} \tag{5-3}$$

where $Pa_{j,t}$, is the percentage of capital amount invested in common stock and cash of mutual fund *j* over the period of ***t***, respectively. $SENT_t$ refers to the PCA sentiment index; *IR* refers to interest rate; *UEMP* refers to unemployment rate, and *GPR* refers to geopolitical risk index. $\phi_j$ refers to the speed of adjustment; $ECT_{j,t-1}$ refers to the error correction term, and $\mu_t$ is group-specific effect.

After defining the linear panel ARDL model, this study uses the nonlinear panel ARDL model proposed by Shin et al. (2014) to determine whether the SENT index has an asymmetric effect on the composition of the mutual fund portfolio. The SENT PCA index is decomposed into its positive and negative partial sums as follows:

$$SENT_t = SENT_0 + SENT_t^{+} + SENT_t^{-} \tag{5-4}$$

where $SENT_t$ is the initial value, and $SENT_t^+$ and $SENT_t^-$ refer to the partical sum processes of positive and negative shocks in PCA sentiment indicators separately.

$$SENT_t^+ = \sum_{k=1}^{t} \Delta SENT_k^+ = \sum_{k=1}^{t} max(\Delta SENT_k, 0) \tag{5-5}$$

$$SENT_t^- = \sum_{k=1}^{t} \Delta SENT_k^- = \sum_{k=1}^{t} min(\Delta SENT_k, 0) \tag{5-6}$$

where *k* refers to the sample unit, and *t* refers to the number of periods.

Incorporating long-run and short-run asymmetry relationships into the linear panel ARDL allows Equation (5-2) to become a nonlinear panel ARDL, Equation (5-7):

$$\begin{aligned} Pa_{j,t} = \alpha_0 + \beta_{0,j} Pa_{j,t-1} + \beta_1 SENT_{t-1}^+ + \beta_2 SENT_{t-1}^- + \beta_3 IR_{t-1} + \beta_4 UEMP_{t-1} + \beta_5 GPR_{t-1} + \\ \sum_{i=1}^{p-1} \lambda_{0,j}\, \Delta Pa_{j,t-i} + \sum_{i=0}^{q-1} \lambda_1\, \Delta SENT_{t-i}^+ + \sum_{i=0}^{q-1} \lambda_2\, \Delta SENT_{t-i}^- + \sum_{i=0}^{q-1} \lambda_3\, \Delta IR_{t-i} + \sum_{i=0}^{q-1} \lambda_4\, \Delta UEMP_{t-i} + \\ \sum_{i=0}^{q-1} \lambda_5\, \Delta GPR_{t-i} + \mu_t + \varepsilon_{j,t} \end{aligned} \tag{5-7}$$

where $\mu_t$ refers to group-specific effect.

As Pesaran et al. (2001) propose, Equation (5-7) can be reformulated as an error correction model:

$$\begin{aligned} \Delta Pa_{j,t} = \phi_j ECT_{j,t-1} + \sum_{i=1}^{p-1} \lambda_{0,j}\, \Delta Pa_{j,t-i} + \sum_{i=0}^{q-1} \lambda_1\, \Delta SENT_{t-i}^+ + \sum_{i=0}^{q-1} \lambda_2\, \Delta SENT_{t-i}^- + \\ \sum_{i=0}^{q-1} \lambda_3\, \Delta IR_{t-i} + \sum_{i=0}^{q-1} \lambda_3\, \Delta UEMP_{t-i} + \sum_{i=0}^{q-1} \lambda_5\, \Delta GPR_{t-i} + \mu_t + \varepsilon_{j,t} \end{aligned} \tag{5-8}$$

Consequently, following the presentation of the empirical method, the empirical findings will be presented in the subsequent section.

## 5.3 Results and Discussion

Before conducting regression, it is necessary to confirm the newly constructed PCA index is not an I(2) variable; therefore, panel unit root tests are applied.

**Table 5–5. Panel unit root test of SENT**

| Method<br>Variables | *IPS* | *Fisher-ADF* | *Fisher-PP* |
|---|---|---|---|
| ***SENT*** | -6. 913* [0.000] | 278.702* [0.000] | 278.702* [0.000] |

Notes: * denotes statistically significant at 5% level. The results demonstrate coefficient and p-value with square brackets.

As shown in Table 5–5, the panel unit root test result of SENT indicates that the PCA index is stationary at I(0) with all the unit root tests. The unit root test results of other control variables are presented in the previous chapter; therefore, the dataset is still in a mixed order of integration, and the panel ARDL is the proper method for this research.

**Table 5–6. The regression results of Panel ARDL with SENT**

| Variables | *Ps* | *Pc* |
|---|---|---|
| | ***Long run equation*** | |
| ***SENT*** | 0.830* | -0.085* |
| | (0.149) | (0.015) |
| ***IR*** | 0.228* | 0.020 |
| | (0.108) | (0.010) |
| ***UEMP*** | 0.170* | -0.024* |
| | (0.074) | (0.008) |
| ***GPR*** | -0.397 | 0.019 |
| | (0.249) | (0.024) |
| | ***Short run equation*** | |
| ***\|COINTEQ*** | -0.346* | -0.588* |
| | (0.023) | (0.027) |
| ***D(SENT)*** | -0.268 | -0.407* |
| | (0.199) | (0.135) |
| ***D(IR)*** | -1.645* | -0.493* |
| | (0.768) | (0.192) |
| ***D(UEMP)*** | -0.198 | -0.136* |
| | (0.104) | (0.039) |
| ***D(GPR)*** | 0. 124 | 0.274* |
| | (0.128) | (0.236) |
| ***C*** | 29. 258* | 0.976* |
| | (2.250) | (0.110) |
| | ***Wald coefficient test*** | |
| *F-statistic* | 11.833* | 13.422* |
| *Chi-square* | 47.331* | 53.688* |

Notes: * denotes statistically significant at 5% level. The results demonstrate coefficient and standard error with parentheses. Ps and Pc refer to the percentage of capital amount of mutual funds invested in common stock, and cash, respectively. SENT means the principal component index, IR means short-term interest rate, UEMP indicates the unemployment rate, and GPR refers to the geopolitical risk index.

The regression results are presented in Table 5–6. The SENT index has a significant positive link with the capital amounts of mutual funds' stock investments, indicating that mutual funds increase their stock investments by 8.3% for every 10 units increase in the SENT index. As for the quantity of cash holdings, the SENT index has a negative effect, suggesting that mutual funds will accumulate more cash while the sentiment index is declining. When investor sentiment increases, mutual funds tend to increase their appetite for risk by investing more in risky assets and less in safe assets. As for control variables, *IR* and *UEMP* have a significant positive impact on stock investment. However, *IR* has a positive influence on cash holdings, whereas *UEMP* has a slight negative impact on mutual fund cash holdings among the control variables.

**Table 5–7. Nonlinear ARDL regression result of SENT**

| Variables | ***Ps*** | ***Pc*** |
|---|---|---|
| | ***Long run equation*** | |
| ***SENT⁺*** | 0.756* | -0.072* |
| | (0.191) | (0.021) |
| ***SENT⁻*** | 0.511* | -0.064* |
| | (0.249) | (0.028) |
| ***IR*** | 0.329* | 0.026* |
| | (0.150) | (0.016) |
| ***UEMP*** | 0.181* | -0.019* |
| | (0.074) | (0.008) |
| ***GPR*** | 0.194 | 0.013 |
| | (0.263) | (0.028) |
| | ***Short run equation*** | |
| ***\|COINTEQ*** | -0.379* | -0.595* |
| | (0.026) | (0.029) |
| ***D(SENT⁺)*** | -0.200 | -0.337 |
| | (0.600) | (0.258) |
| ***D(SENT⁻)*** | -0.286 | -0.437* |
| | (0.167) | (0.193) |
| ***D(IR)*** | -1.846* | -0.488* |
| | (0.737) | (0.233) |
| ***D(UEMP)*** | -0.229* | -0.139* |
| | (0.113) | (0.038) |
| ***D(GPR)*** | 0.028 | 0.275* |
| | (0.125) | (0.234) |
| ***C*** | 30.732* | 0.966* |
| | (2.447) | (0.123) |

Notes: * denotes statistically significant at 5% level. The results demonstrate coefficient and standard error with parentheses. + indicates positive change, - indicates negative change, Ps and Pc refer to the percentage of capital amount of mutual funds invested in common stock, and cash, respectively. SENT means the principal component index, + indicates positive change, - indicates negative change, IR means short-term interest rate, UEMP indicates the unemployment rate, and GPR refers to the geopolitical risk index. COINTEQ refers to error correction term.

Table 5-7 reveals the results of the nonlinear panel ARDL regression of the SENT index. As specified in Equations (5-6)(5-6) and (5-7), the SENT index is split into positive and negative shocks. I find that, in the long run, positive shocks and negative shocks to investor sentiment have significantly positive influences on mutual funds' percentage of stock investment. The positive shocks have larger coefficient than negative shocks. As for the cash holdings, they are affected by investor sentiment fluctuations in the long run. Mutual funds will decrease their cash holdings during positive sentiment shocks increasing and increase the cash amount when negative shocks are declining. In the short run, only the negative shocks of SENT have a significant influence on mutual funds' stock investments and cash amounts, but the significance of positive shocks is absent for either stock or cash.

The results indicate that mutual fund portfolio composition is more sensitive toward the positive shocks of investor sentiment and is consistent with the results of other sentiment proxies. Therefore, mutual funds act against the value function, and they have a higher subjective valuation to good news or positive sentiment than to negative sentiment.

**Table 5–8. Wald test of asymmetry effect**

| | ***Ps*** | ***Decision*** | ***Pc*** | ***Decision*** |
|---|---|---|---|---|
| ***$SENT_{LR}$*** | 9.75 [0.002] | Yes | 0.86 [0.358] | No |
| ***$SENT_{SR}$*** | 0.02 [0.902] | No | 0.09 [0.763] | No |

* Ps and Pc refer to the percentage of the capital amount of mutual funds invested in common stock, and cash, respectively. SENT means the principal component index. LR and SR represent long run and short run asymmetry. The results show Chin-square and p-value in bracket.

The Wald test is conducted to investigate the asymmetry effect. As illustrated in Table 5–8, the null hypothesis of $H_0 = \beta^+ = \beta^-$ is rejected for *Ps* regressions in the long run, which means that the SENT index has a long-run asymmetric effect on the percentage of mutual funds' portfolios invested in stocks. However, the null hypothesis cannot be rejected in the short-run *Ps* regression or the long run or short run of cash holdings, which means that the asymmetric effect is not present on the short-run relationship between sentiment and mutual fund stock investment, and there is no asymmetric effect between sentiment and mutual fund cash holdings in both the long run and short run.

# 5.4 Robustness Checks

This section rigorously evaluates the robustness of the results through the implementation of three distinct methodological approaches. Initially, it utilizes the Wald test to assess the joint significance of the variables involved. Subsequently, the analysis incorporates the Dumitrescu-Hurlin panel causality test(Dumitrescu and Hurlin, 2012), offering a nuanced examination of causality within the panel data framework. Finally, the robustness is further substantiated through conducting regressions on sub-samples of the data, ensuring the consistency and reliability of the findings across different time periods.

**Table 5–9. Wald test of SENT index**

| | ***Ps*** | | ***Pc*** | |
|---|---|---|---|---|
| ***SENT*** | *F-statistic* | *Chi-square* | *F-statistic* | *Chi-square* |
| | 11.833* | 47.331* | 13.422* | 53.688* |

Notes: * denotes Significant at 5% level. Ps and Pc refer to the percentage of capital amount of mutual funds invested in common stock, and cash, respectively.

| | ***Ps*** | ***Decision*** | ***Pc*** | ***Decision*** |
|---|---|---|---|---|
| ***$SENT_{LR}$*** | 9.75 [0.002] | Yes | 0.85 [0.358] | No |
| ***$SENT_{SR}$*** | 0.02 [0.902] | No | 0.09 [0.763] | No |

Ps and Pc refer to the percentage of capital amount of mutual funds invested in common stock, and cash, respectively. The results show Chin-square and p-value in bracket.

The Wald test results presented in Table 5–9 reveal the jointly significant explanatory variables in the linear panel ARDL regressions, where $H_0: \beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$ is the null hypothesis. If the null hypothesis is rejected, all variables are crucial to include.

The Wald test shows significance for F-statistics and Chi-squared of both stock percentage and cash holdings percentage regressions. This indicates that the null hypothesis is rejected, and the alternative hypothesis is accepted, which means that all estimated coefficients in the models are efficient and important to the models.

**Table 5–10. Dumitrescu-Hurlin causality test of linear regression**

| | ***Ps*** | | | ***Pc*** | |
|---|---|---|---|---|---|
| | Wald-statistics | p-value | | Wald-statistics | p-value |
| *SENT does not homogeneously cause Ps* | 10.223 | 0.000 | *SENT does not homogeneously cause Pc* | 7.440 | 0.000 |
| *Ps does not homogeneously cause SENT* | 18.081 | 0.000 | *Pc does not homogeneously cause SENT* | 14.638 | 0.000 |
| *$SENT^+$ does not homogeneously cause Ps* | 7.930 | 0.000 | *$SENT^+$ does not homogeneously cause Pc* | 5.727 | 0.000 |
| *Ps does not homogeneously cause $SENT^+$* | 3.703 | 0.000 | *Pc does not homogeneously cause $SENT^+$* | 3.525 | 0.000 |
| *$SENT^-$ does not homogeneously cause Ps* | 4.377 | 0.000 | *$SENT^-$ does not homogeneously cause Pc* | 6.577 | 0.000 |
| *Ps does not homogeneously cause $SENT^-$* | 7.723 | 0.000 | *Pc does not homogeneously cause $SENT^-$* | 7.453 | 0.000 |

Ps and Pc refer to the percentage of capital amount of mutual funds invested in common stock, and cash, respectively. SENT means the principal component index, + indicates positive change, - indicates negative change.

The Dumitrescu-Hurlin (Dumitrescu and Hurlin, 2012) panel causality between concern variables and mutual funds' stock investments and cash holdings are illustrated in Table 5–10. Thus, in the forthcoming analysis, I aim to persistently encourage the reader to view the estimated causal effects as suggestive, rather than definitive. The results signify the bidirectional causality between the PCA SENT index and dependent variables.

To further test the robustness of the results, I employed a sub-sample period from 2010 Q1 to 2019 Q4, which excludes serious crises, and investor sentiment is relatively stable.

**Table 5–11. Robustness check of sub-sample**

| | *Ps* | *Pc* |
|---|---|---|
| ***Long run equation*** | | |
| ***SENT*** | 0.830* <br> (0.330) | 0.040 <br> (0.028) |
| ***IR*** | 0.246 <br> (0.154) | 0.018 <br> (0.010) |
| ***UEMP*** | 0.108 <br> (0.133) | 0.022 <br> (0.008) |
| ***GPR*** | 0.344 <br> (0.257) | 0.049* <br> (0.024) |
| ***Short run equation*** | | |
| ***\|COINTEQ*** | -0.341* <br> (0.027) | -0.606* <br> (0.030) |
| ***D(SENT)*** | -0.295 <br> (0.214) | -0.498* <br> (0.209) |
| ***D(IR)*** | -0.405 <br> (0.472) | -1.281* <br> (0.489) |
| ***D(UEMP)*** | 0.224 <br> (0.352) | -0.164 <br> (0.342) |
| ***D(GPR)*** | -0. 016 <br> (0.111) | 0.218* <br> (0.100) |
| ***C*** | 29. 025* <br> (2.571) | 0.743* <br> (0.125) |

Notes: * denotes statistically significant at 5% level. The results demonstrate coefficient and standard error with parentheses. Ps and Pc refer to the percentage of capital amount of mutual funds invested in common stock, and cash, respectively. SENT means the principal component index, + indicates positive change, - indicates negative change, IR means short-term interest rate, UEMP indicates the unemployment rate, and GPR refers to the geopolitical risk index. COINTEQ refers to error correction term.

The results in Table 5–11 indicates that the SENT index is still volatile and has a significant positive relationship with stock investment amounts in mutual funds in the long run. However, compared to the full sample result, the SENT index does not impact mutual fund cash holdings in the long run but still has a significantly negative impact in the short run. It is understandable that mutual funds may have limited cash movement during stable times in the long run. However, they still react to short-term sentiment shocks. This may result from mutual funds viewing short-term sentiment shocks as a risk signal. Then, mutual funds try to protect their exposure by adjusting safe asset amounts to avoid sudden market fluctuation losses. Therefore, these results are consistent with previous results overall, and their robustness is confirmed.

**Table 5–12. Robustness check of nonlinear sub-sample**

| Variables | *Ps* | *Pc* |
|---|---|---|
| | ***Long run equation*** | |
| ***SENT⁺*** | 0.128 | 0.047 |
| | (0.387) | (0.036) |
| ***SENT⁻*** | 1.504* | -0.012 |
| | (0.443) | (0.043) |
| ***IR*** | -0.010 | 0.029 |
| | (0.169) | (0.016) |
| ***UEMP*** | -0.717* | 0.052* |
| | (0.280) | (0.026) |
| ***GPR*** | -0.268 | 0.074* |
| | (0.320) | (0.030) |
| | ***Short run equation*** | |
| ***\|COINTEQ*** | -0.345* | -0.612* |
| | (0.028) | (0.030) |
| ***D(SENT⁺)*** | -0.570 | -0.312 |
| | (0.412) | (0.297) |
| ***D(SENT⁻)*** | 0.489 | -0.423 |
| | (0.462) | (0.319) |
| ***D(IR)*** | -0.453 | -1.174* |
| | (0.472) | (0.474) |
| ***D(UEMP)*** | 0.173 | -0.137 |
| | (0.335) | (0.342) |
| ***D(GPR)*** | 0.120 | 0.209* |
| | (0.108) | (0.103) |
| ***C*** | 34.627* | 0.390* |
| | (3.003) | (0.124) |

Notes: * denotes statistically significant at 5% level. The results demonstrate coefficient and standard error with parentheses. + indicates positive change, - indicates negative change, Ps and Pc refer to the percentage of capital amount of mutual funds invested in common stock, and cash, respectively. SENT means the principal component index, + indicates positive change, - indicates negative change, IR means short-term interest rate, UEMP indicates the unemployment rate, and GPR refers to the geopolitical risk index. COINTEQ refers to error correction term.

In addition, as Table 5–12 demonstrates, the negative shocks of SENT have significant influence on stock investment of mutual funds in the long run, while the positive shocks do not. This suggests that mutual funds may react to negative shocks of sentiment (loss) more than positive shocks (gains) during stable period. The value function of prospect theory could be applied to explain the investment activities and portfolio composition adjustment of mutual funds. As for the cash holdings, there is no significant relationship from sentiment fluctuation in both the long run and short run. The results are consistent with Chapter 4; mutual funds' activities fit into loss aversion only when the market is relatively stable and without severe crisis but will react more to positive sentiment changes when the crisis periods are included. As mentioned, in periods of crisis, investors may demonstrate an increased propensity to seek and respond favourably to positive information, particularly when circumstances appear dire. Conversely, in periods of stability, mutual funds often exhibit characteristics of loss aversion, responding more significantly to negative information or adverse shifts in sentiment. Overall, these findings indicate that prospect theory is a valuable framework for comprehending the behaviours of institutional investors, such as those managing mutual funds.

## 5.5 Conclusion

This chapter builds a new investor sentiment index calibrated for capturing the mutual fund portfolio composition in the US market by aggregating the information contained in five widely used and multiple-based sentiment proxies using the PCA method. I find that this new investor sentiment index has higher correlation with mutual fund portfolio composition compared to the conventional BW Index. The results demonstrate that mutual funds will increase their proportion of risky investments when sentiment increases.

Importantly, I find that institutional investors like mutual funds, especially their stock investments, may react to fluctuation of sentiment asymmetrically. Mutual funds may exhibit varying responses to shifts in investor sentiment during stable and crisis periods. During crises, these funds appear more responsive to rising sentiment than to declining sentiment, a trend that contradicts the value function and concept of loss aversion. This pattern may arise because the market conditions are perceived as having deteriorated to their nadir, leading investors to become desensitized to further negative information or sentiment. As a result, they display a heightened interest in and responsiveness to positive news or sentiment. Conversely, when crisis periods are excluded from analysis, mutual funds show increased sensitivity to negative sentiment. This behaviour aligns with the value function principle, suggesting that investors might place greater emphasis on losses than gains. Therefore, while the value function and loss aversion principles of prospect theory can elucidate the portfolio strategies of mutual funds, their applicability is limited to non-crisis periods.

# Chapter 6. General Conclusion

This thesis aims to explore the relationship between multiple investor sentiment proxies—such as surveys, market indicators, and internet activity—and the changing portfolio composition of mutual funds in the U.S. market over time. Additionally, it seeks to identify correlations between different investor sentiment indices and mutual fund allocations across various asset categories while investigating potential asymmetric effects. Another objective is to develop a new composite sentiment index that reflects mutual fund portfolio composition, incorporating different sentiment indicators.

Chapter 3 focuses specifically on the relationship between various investor sentiment measures and mutual fund portfolio allocation, particularly in stocks and cash holdings. The study examines this relationship using multiple sentiment indicators, including survey-based measures (e.g., Consumer Confidence Index), internet-based metrics (e.g., Google search volumes for terms like "financial crisis" and "economic recession"), and market-based indicators (e.g., the number of IPOs, first-day IPO returns, and the Volatility Index), along with the composite Baker and Wurgler (BW) Index.

The results indicate that all of the investor sentiment measures significantly influence mutual fund portfolio composition. This suggests that mutual fund managers adjust their portfolios in response to changes in investor sentiment. Specifically, consumer confidence and Google search volumes as proxies for sentiment impact both stock allocations and cash holdings within mutual funds. This highlights the influence of broader economic sentiment, as captured through surveys and internet activity, in shaping institutional investment decisions. Such findings consistent with the finding of Tetlock (2007) and Da et al. (2015) that mutual funds do not operate in isolation but are influenced by the sentiments and expectations of the public. Interestingly, the composite BW Index, which integrates multiple market-based indicators, was found to have the highest correlation with mutual fund portfolio composition compared to single indicators.

Chapter 4 examines the nonlinear and asymmetric relationships between investor sentiment and mutual fund portfolio composition. The findings indicate that changes in sentiment, whether positive or negative, affect the overall structure of mutual fund portfolios, with an asymmetry in how these sentiment indices impact stock holdings versus cash allocations. The asymmetric effects observed—where mutual funds react differently to positive and negative sentiment—can be interpreted through prospect theory, particularly the concepts of the value function and loss aversion (Kahneman and Tversky, 1979). According to prospect theory, investors are more sensitive to losses than to gains, which should theoretically lead to a stronger reaction to negative sentiment. However, the findings suggest that during periods of crisis, mutual funds may exhibit a stronger response to positive sentiment, potentially due to a perception that the market has bottomed out and there is more to gain from an expected recovery.

This deviation from expected behaviour, as suggested by the value function, implies that institutional investors might temporarily override loss aversion in favour of capitalizing on potential gains when market conditions are particularly dire. These findings contribute to the broader understanding of investor sentiment and behavioural biases by illustrating how institutional investors, such as mutual funds, might adjust their strategies in response to market conditions and sentiment changes.

The asymmetric response pattern indicates that during stable periods, mutual funds align more closely with loss aversion principles, while during crises, they might shift focus towards potential gains.

In Chapter 5, the final empirical chapter, the study constructs a composite index using Principal Component Analysis (PCA) by incorporating multiple sentiment indicators, such as the Consumer Confidence Index, Consumer Sentiment Index, Google search volumes, the number of new IPOs, and the Volatility Index. The composite sentiment index performs well in reflecting mutual fund portfolio composition. However, the Baker and Wurgler (BW) Index is only constructed using market-based proxies, which may omit important information. Therefore, it is reasonable to consider survey-based and internet-based sentiment measurements, as they also have the potential to impact mutual fund portfolios, as demonstrated in the previous chapters, rather than relying solely on market-based indices. This research follows the method by Baker and Wurgler (2006), building a first-stage component index with current and lagged values of five sentiment indicators, and then decide whether to use the current or lagged values by selecting which have higher correlations with the first-stage index. Given the results, the new PCA index has better performance in reflecting the mutual fund portfolio allocation change in both the capital percentage of common stock investment and cash holdings compared to the BW Index.

Crucially, this research identifies those institutional investors, such as mutual funds, particularly in their equity investments, may exhibit asymmetric reactions to changes in sentiment. Mutual funds' responsiveness to investor sentiment can vary markedly between stable times and periods of crisis. In times of crisis, mutual funds are observed to be more reactive to increasing sentiment compared to decreasing sentiment. This response pattern appears to be at odds with the expected behaviours suggested by the value function and loss aversion concept. One explanation for this could be that market conditions are perceived to be at their lowest point, leading to diminished sensitivity to further negative news or sentiment among investors. Consequently, there is an increased focus on and reaction to positive developments and sentiment. In contrast, outside of crisis periods, mutual funds are more prone to react to negative sentiment, aligning more closely with the value function principle. This indicates a tendency for investors to prioritize losses over gains. Thus, while prospect theory's principles of the value function and loss aversion can shed light on the strategic approaches of mutual funds in portfolio management, their relevance is notably more pronounced during periods devoid of crisis.

Overall, this thesis examines the impact of investor sentiment on portfolio composition. The research confirms that investor sentiment significantly influences the allocation of assets within mutual fund portfolios. The study's findings demonstrate that both positive and negative shifts in sentiment have distinct and asymmetric effects on portfolio structure, particularly in stock investments and cash holdings. This reinforces the importance of considering behavioural biases, such as those outlined in prospect theory, when analysing investment decisions.

In addition, this thesis finds asymmetric effects of investor sentiment on portfolio composition and highlights the behavioural biases of mutual funds, where funds are more reactive to positive sentiment during crises and to negative sentiment during stable periods. This aligns with the principles of the value function in prospect theory, where investors' responses to gains and losses are unequal. This thesis provides empirical evidence supporting the application of prospect theory in understanding investor behaviour in portfolio management. It also suggests that traditional models may need to be adjusted to better account for behavioural factors, particularly in volatile market conditions.

## 6.1 Contributions and Implications

This thesis makes several significant contributions to the fields of behavioural finance, portfolio management, and the study of the relationship between investor sentiment and mutual fund portfolio composition. This research advances our understanding of how different investor sentiment proxies—such as survey-based indicators, market-based metrics, and internet-derived sentiment—affect mutual fund portfolio allocation over time. The study fills a gap in the literature by demonstrating the impact of sentiment on the composition of mutual fund portfolios, emphasizing the role of asymmetric effects and behavioural biases. Next, A novel composite sentiment index was constructed using Principal Component Analysis (PCA), incorporating diverse sentiment indicators from surveys, internet search data, and market-based metrics. This index provides a more comprehensive measure of investor sentiment, addressing limitations in existing models that primarily focus on market-based proxies. The new index shows superior performance in reflecting mutual fund portfolio changes, offering a robust tool for future research and practical applications.
At last, this research underscores the importance of incorporating behavioural biases into portfolio management strategies. By highlighting how investor sentiment influences mutual fund portfolio composition, particularly in times of crisis, the study contributes to the development of more resilient investment strategies that account for investor psychology.

The research findings have significant implications for both institutional investors and individual investors, particularly in the context of portfolio management and investment decision-making. For individual investor groups, the implications as following:

Investor Sentiment and Decision-Making: Individual investors often rely on sentiment obtained from the broader market to guide their investment choices. By understanding the role of sentiment, individual investors can better interpret market signals and avoid common pitfalls such as buying into overvalued markets during periods of high optimism or selling off assets during market panics. This knowledge empowers individual investors to adopt a more contrarian approach, where they might buy when sentiment is low and sell when sentiment is high, thereby potentially improving their investment outcomes.

Asymmetric Effects on Portfolio Construction: The research highlights how market reactions to sentiment are often asymmetric, with negative sentiment having a more pronounced impact on asset prices than positive sentiment. Individual investors can use this insight to structure their portfolios in a way that mitigates downside risk. For example, during periods of anticipated negative sentiment, they might increase their holdings in safe assets or allocate more to cash to preserve capital. Conversely, in more stable or optimistic markets, they might take on more risk by increasing their exposure to equities or other growth-oriented investments.

Behavioural Biases and Investment Strategies: Individual investors are particularly susceptible to behavioural biases such as loss aversion, overconfidence, and herding. The research underscores the importance of recognizing these biases and taking steps to mitigate their impact. For instance, by understanding their own tendencies towards loss aversion, individual investors can set more rational stop-loss limits and avoid panic selling during market downturns.

Incorporating Behavioural Insights into Decision-Making: The application of prospect theory and behavioural finance principles can further enhance individual investment strategies. By recognizing that losses often loom larger than gains, individual investors can develop more balanced and risk-aware portfolios. This might involve setting realistic expectations for returns, diversifying across asset classes to manage risk.

Then, the implication for institutional investors is demonstrated as follows:

Sentiment-Driven Strategies: Institutional investors can leverage the findings on investor sentiment to refine their investment strategies. For instance, by monitoring sentiment indicators, institutional investors can anticipate market trends and adjust their asset allocations accordingly. In periods of high positive sentiment, they might reduce exposure to overvalued assets and increase holdings in underappreciated sectors. Conversely, during periods of negative sentiment, they might capitalize on market dislocations by acquiring undervalued assets.

Managing Asymmetric Risk: The research highlights the asymmetric effects of sentiment on asset prices, particularly the heightened impact of negative sentiment. Institutional investors can use this knowledge to enhance their risk management practices. For example, they might implement more robust hedging strategies during periods of negative sentiment or increase their focus on downside protection. This could involve the use of options, futures, or other derivative instruments to manage risk and protect portfolio value during market downturns.

Addressing Behavioural Biases in Investment Committees: Institutional investment decisions are often made by committees, where behavioural biases can still play a significant role. The research findings suggest that even sophisticated investors are not immune to biases such as herding and overconfidence. By incorporating behavioural finance principles into their decision-making frameworks, institutional investors can develop more disciplined investment processes. This might include setting clear criteria for investment decisions, using quantitative models to reduce subjective judgment, and conducting regular reviews to ensure that decisions are not unduly influenced by market sentiment or emotional reactions.

Incorporating Prospect Theory in Portfolio Management: Institutional investors can benefit from applying the principles of prospect theory to their portfolio management practices. Understanding that losses are perceived more strongly than gains can inform how risk is managed within portfolios. For example, during periods of market stability, institutional investors might adopt a more aggressive stance, aligning with the typical value function that prioritizes potential gains. However, during periods of market stress, a more conservative approach might be warranted, focusing on capital preservation and minimizing losses.

# 6.2 Future Research

One of the primary limitations of this research lies in the constraints related to data availability and frequency. While the study employed quarterly data from the CRSP database and other sentiment indicators, the use of quarterly data may not fully capture the nuances and short-term fluctuations in investor sentiment and mutual fund portfolio composition. The quarterly aggregation of data could result in the loss of details that might be observable with higher frequency data, such as monthly observations. This limitation could impact the precision of the results and may lead to oversights in understanding the short-term dynamics between investor sentiment and portfolio allocation. The reliance on quarterly data limits the ability to observe more immediate reactions of mutual funds to shifts in sentiment, particularly during volatile periods when sentiment can change rapidly. This constraint could affect the interpretation of the study's findings, particularly in understanding the full scope of sentiment’s impact on portfolio management in real-time. Future studies could explore the use of higher frequency data, such as monthly or daily data, to provide a more granular analysis of the relationship between investor sentiment and mutual fund portfolio composition. Additionally, accessing alternative datasets that offer more detailed insights into investor behaviour could help overcome the limitations imposed by quarterly data.

The research relies on several methodological assumptions, particularly in the construction of the composite sentiment index using Principal Component Analysis (PCA) and the application of regression models to analyse the impact of sentiment on portfolio composition. While these methodologies are robust and widely accepted, they come with inherent limitations. Firstly, the use of PCA to construct the composite sentiment index assumes that the underlying sentiment indicators are linearly related and that the principal components adequately capture the variance in the data. However, this assumption might not hold true if the relationship between sentiment indicators is more complex or nonlinear, potentially leading to an oversimplification of the sentiment dynamics. Secondly, the regression models applied in the study assume linearity and homoscedasticity, which may not fully capture the potential nonlinear or asymmetric relationships between sentiment and portfolio composition. These assumptions could limit the models' ability to detect more nuanced effects or interactions within the data. Future research could explore alternative methodological approaches that account for potential nonlinearities and interactions in the data. For instance, machine learning techniques or nonlinear modelling could be employed to capture more complex relationships between sentiment and portfolio composition. Additionally, validating the PCA findings with other dimensionality reduction techniques could enhance the robustness of the composite sentiment index.

In addition, this research incorporates prospect theory into mutual funds portfolio composition, but only a component of prospect theory, the value function. The remainder of prospect theory, like the weighting function, is an interesting area to investigate, such as whether mutual funds are risk-seeking or risk-adverse during a specific time. Furthermore, these investor sentiment indices, including the existing or newly constructed indices, are only used to determine their effects on the portfolio composition, especially the institutional portfolio, in this study. It is necessary to consider how they influence the individual portfolio composition of retail investors if the data is available. Furthermore, it is interesting to investigate how these investor sentiment indicators reveal the returns of selected portfolios, no matter the individual and institutional perspectives. At last, this study focuses on institutional investors, particularly mutual funds, within the US market. The findings may not be directly applicable to other markets. Further research is needed to explore the effects of sentiment on different investor groups and in different countries.

## 6.3 Policy Recommendations

Based on the findings of this research, which explores the influence of multiple investor sentiment proxies on mutual fund portfolio composition and the application of the value function from prospect theory. The following policy recommendations are proposed:

This research underscores the importance of integrating insights from investor sentiment into broader portfolio management and risk management practices for both individual and institutional investors.

Dynamic Asset Allocation: Both individual and institutional investors can benefit from adopting dynamic asset allocation strategies that adjust based on prevailing sentiment and market conditions. This involves regularly rebalancing portfolios to align with changing sentiment indicators and adjusting exposure to different asset classes as market conditions evolve.

Enhanced Risk Management Frameworks: The findings suggest that traditional risk management frameworks might need to be augmented with sentiment analysis and behavioural insights. This could involve incorporating sentiment indicators into risk models and developing contingency plans for periods of extreme market sentiment.

Investment Strategy Innovation: Finally, the research points to the potential for innovation in investment strategies that explicitly account for sentiment and behavioural factors. For example, the development of sentiment-driven trading algorithms or the creation of investment products that hedge against behavioural biases could offer new opportunities for both individual and institutional investors.

Incorporating an understanding of investor sentiment into investment strategies, portfolio construction, and risk management practices can significantly enhance decision-making for both individual and institutional investors. By leveraging insights from prospect theory, investors can develop more resilient and adaptive strategies that are better equipped to navigate the complexities of financial markets. These approaches not only improve investment outcomes but also contribute to a more sophisticated and informed investment process that accounts for the psychological factors influencing market behaviour.

# References


Abdi, H., and Williams, L.J. (2010). Principal component analysis. *WIREs Computational Statistics* 2(4), 433-459. doi: https://doi.org/10.1002/wics.101.

Akbas, F., Armstrong, W.J., Sorescu, S., and Subrahmanyam, A. (2015). Smart money, dumb money, and capital market anomalies. *Journal of Financial Economics* 118(2), 355-382. doi: https://doi.org/10.1016/j.jfineco.2015.07.003.

Amihud, Y. (2002). Illiquidity and stock returns: cross-section and time-series effects. *Journal of Financial Markets* 5(1), 31-56. doi: https://doi.org/10.1016/S1386-4181(01)00024-6.

Amihud, Y., and Mendelson, H. (1986). Liquidity and Stock Returns. *Financial Analysts Journal* 42(3), 43-48. doi: 10.2469/faj.v42.n3.43.

Antweiler, W., and Frank, M.Z. (2004). Is All That Talk Just Noise? The Information Content of Internet Stock Message Boards. *The Journal of Finance* 59(3), 1259-1294. doi: https://doi.org/10.1111/j.1540-6261.2004.00662.x.

Atanasov, V. (2021). Unemployment and aggregate stock returns. *Journal of Banking & Finance* 129, 106159. doi: https://doi.org/10.1016/j.jbankfin.2021.106159.

Baker, M., and Stein, J.C. (2004). Market liquidity as a sentiment indicator. *Journal of Financial Markets* 7(3), 271-299. doi: https://doi.org/10.1016/j.finmar.2003.11.005.

Baker, M., and Wurgler, J. (2000). The Equity Share in New Issues and Aggregate Stock Returns. *The Journal of Finance* 55(5), 2219-2257. doi: https://doi.org/10.1111/0022-1082.00285.

Baker, M., and Wurgler, J. (2006). Investor Sentiment and the Cross-Section of Stock Returns. *The Journal of Finance* 61(4), 1645-1680. doi: https://doi.org/10.1111/j.1540-6261.2006.00885.x.

Baker, M., and Wurgler, J. (2007). Investor Sentiment in the Stock Market. *Journal of Economic Perspectives* 21(2), 129-152. doi: 10.1257/jep.21.2.129.

Barberis, N., Shleifer, A., and Vishny, R. (1998). A model of investor sentiment1We are grateful to the NSF for financial support, and to Oliver Blanchard, Alon Brav, John Campbell (a referee), John Cochrane, Edward Glaeser, J.B. Heaton, Danny Kahneman, David Laibson, Owen Lamont, Drazen Prelec, Jay Ritter (a referee), Ken Singleton, Dick Thaler, an anonymous referee, and the editor, Bill Schwert, for comments.1. *Journal of Financial Economics* 49(3), 307-343. doi: https://doi.org/10.1016/S0304-405X(98)00027-0.

Barberis, N.C. (2013). Thirty Years of Prospect Theory in Economics: A Review and Assessment. *Journal of Economic Perspectives* 27(1), 173-196. doi: 10.1257/jep.27.1.173.

Basu, D., and Chawla, D. (2010). An Empirical Test of CAPM—The Case of Indian Stock Market. *Global Business Review* 11(2), 209-220. doi: 10.1177/097215091001100206.

Beer, F., Hamdi, B., and Zouaoui, M. (2018). Investors' sentiment and accruals anomaly: European evidence. *Journal of Applied Accounting Research* 19(4), 500-517. doi: 10.1108/JAAR-03-2017-0043.

Beer, F., Hervé, F., and Zouaoui, M. (2012). "Is big brother watching us? Google, investor sentiment and the stock market", in: *Economics Bulletin, Forthcoming.*).

Bekiros, S., Gupta, R., and Kyei, C. (2016). A non-linear approach for predicting stock returns and volatility with the use of investor sentiment indices. *Applied Economics* 48(31), 2895-2898. doi: 10.1080/00036846.2015.1130793.

Bekiros, S., Jlassi, M., Lucey, B., Naoui, K., and Uddin, G.S. (2017). Herding behavior, market sentiment and volatility: Will the bubble resume? *The North American Journal of Economics and Finance* 42, 107-131. doi: https://doi.org/10.1016/j.najef.2017.07.005.

Ben-Rephael, A., Kandel, S., and Wohl, A. (2012). Measuring investor sentiment with mutual fund flows. *Journal of Financial Economics* 104(2), 363-382. doi: https://doi.org/10.1016/j.jfineco.2010.08.018.

Bergman, N.a.R., S (2008). Investor Sentiment and Corporate Disclosure. *Journal of Accounting Research* 46(5), 1057-1083. doi: https://doi.org/10.1111/j.1475-679X.2008.00305.x.

Berk, J.B., and Stanton, R. (2007). Managerial Ability, Compensation, and the Closed-End Fund Discount. *The Journal of Finance* 62(2), 529-556. doi: https://doi.org/10.1111/j.1540-6261.2007.01216.x.

Bilel, H., and Mondher, K. (2021). What Can explain catering of dividend? Environment information and investor sentiment. *Journal of Economics and Finance* 45(3), 428-450. doi: 10.1007/s12197-021-09540-0.

Black, F. (1986). Noise. *The Journal of Finance* 41(3), 528-543. doi: https://doi.org/10.1111/j.1540-6261.1986.tb04513.x.

Bodnaruk, A., and Simonov, A. (2016). Loss-Averse Preferences, Performance, and Career Success of Institutional Investors. *The Review of Financial Studies* 29(11), 3140-3176. doi: 10.1093/rfs/hhw053.

Bollen, J., Mao, H., and Pepe, A. (2021). Modeling Public Mood and Emotion: Twitter Sentiment and Socio-Economic Phenomena. *Proceedings of the International AAAI Conference on Web and Social Media* 5(1), 450-453. doi: 10.1609/icwsm.v5i1.14171.

Bollen, J., Mao, H., and Zeng, X. (2011). Twitter mood predicts the stock market. *Journal of Computational Science* 2(1), 1-8. doi: https://doi.org/10.1016/j.jocs.2010.12.007.

Brennan, M.J., Chordia, T., Subrahmanyam, A., and Tong, Q. (2012). Sell-order liquidity and the cross-section of expected stock returns. *Journal of Financial Economics* 105(3), 523-541. doi: https://doi.org/10.1016/j.jfineco.2012.04.006.

Brennan, M.J., and Subrahmanyam, A. (1996). Market microstructure and asset pricing: On the compensation for illiquidity in stock returns. *Journal of Financial Economics* 41(3), 441-464. doi: https://doi.org/10.1016/0304-405X(95)00870-K.

Brown, G.W., and Cliff, M.T. (2004). Investor sentiment and the near-term stock market. *Journal of Empirical Finance* 11(1), 1-27. doi: https://doi.org/10.1016/j.jempfin.2002.12.001.

Brown, Gregory W., and Cliff, Michael T. (2005). Investor Sentiment and Asset Valuation. *The Journal of Business* 78(2), 405-440. doi: 10.1086/427633.

Brown, S.J., Goetzmann, W.N., Hiraki, T., and Shiraishi, N. (2003). An analysis of the relative performance of Japanese and foreign money management. *Pacific-Basin Finance Journal* 11(4), 393-412. doi: https://doi.org/10.1016/S0927-538X(03)00046-5.

Bu, Q. (2021). Are All the Sentiment Measures the Same? *Journal of Behavioral Finance*, 1-10. doi: 10.1080/15427560.2021.1949718.

Bu, Q., and Forrest, J. (2021). Comparing sentiment measures in mutual fund performance. *International Journal of Managerial Finance* 17(3), 478-493. doi: 10.1108/IJMF-02-2020-0092.

Caparrelli, F., D'Arcangelis, A.M., and Cassuto, A. (2004). Herding in the Italian Stock Market: A Case of Behavioral Finance. *Journal of Behavioral Finance* 5(4), 222-230. doi: 10.1207/s15427579jpfm0504_5.

Chakraborty, M., and Subramaniam, S. (2020). Asymmetric relationship of investor sentiment with stock return and volatility: evidence from India. *Review of Behavioral Finance* 12(4), 435-454. doi: 10.1108/RBF-07-2019-0094.

Chakravarty, S. (2001). Stealth-trading: Which traders' trades move stock prices? *Journal of Financial Economics* 61(2), 289-307. doi: https://doi.org/10.1016/S0304-405X(01)00063-0.

Chang, Y.Y.C., Faff, R.W., and Hwang, C.-Y. (2012). "Local and global sentiment effects, and the role of legal, information and trading environments", in: *Information and Trading Environments (February 28, 2012).*).

Chau, F., Deesomsak, R., and Koutmos, D. (2016). Does investor sentiment really matter? *International Review of Financial Analysis* 48, 221-232. doi: https://doi.org/10.1016/j.irfa.2016.10.003.

Che-Yahya, N., Abdul-Rahim, R., and Yong, O. (2014). Influence of institutional investors' participation on flipping activity of Malaysian IPOs. *Economic Systems* 38(4), 470-486. doi: https://doi.org/10.1016/j.ecosys.2014.03.002.

Chiu, H.H., and Kini, O. (2014). Equity Issuances, Equity Mutual Fund Flows, and Noise Trader Sentiment*. *Review of Finance* 18(2), 749-802. doi: 10.1093/rof/rft009.

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

Choi, J.J., Laibson, D., and Metrick, A. (2002). How does the Internet affect trading? Evidence from investor behavior in 401(k) plans. *Journal of Financial Economics* 64(3), 397-421. doi: https://doi.org/10.1016/S0304-405X(02)00130-7.

Choi, N.Y., and Sias, R.W. (2012). Why Does Financial Strength Forecast Stock Returns? Evidence from Subsequent Demand by Institutional Investors. *The*

*Review of Financial Studies* 25(5), 1550-1587. doi: 10.1093/rfs/hhs001.
Chue, T.K., and Mian, G.M. (2022). Investor sentiment and mutual fund stock picking. *Applied Economics Letters* 29(17), 1620-1625. doi: 10.1080/13504851.2021.1951440.
Ciminelli, G., Rogers, J., and Wu, W. (2022). The effects of U.S. monetary policy on international mutual fund investment. *Journal of International Money and Finance* 127, 102676. doi: https://doi.org/10.1016/j.jimonfin.2022.102676.
Clark, J.M. (1918). Economics and Modern Psycholoy: I. *Journal of Political Economy* 26(1), 1-30. doi: 10.1086/253060.
Cornell, B., Landsman, W.R., and Stubben, S. (2011). Do Institutional Investors and Security Analysts Mitigate the Effects of Investor Sentiment? *Available at SSRN 1848207*.
Cornelli, F., Goldreich, D., and Ljungqvist, A. (2006). Investor Sentiment and Pre-IPO Markets. *The Journal of Finance* 61(3), 1187-1216. doi: https://doi.org/10.1111/j.1540-6261.2006.00870.x.
Corredor, P., Ferrer, E., and Santamaria, R. (2013). Investor sentiment effect in stock markets: Stock characteristics or country-specific factors? *International Review of Economics & Finance* 27, 572-591. doi: https://doi.org/10.1016/j.iref.2013.02.001.
Curtis, G. (2004). Modern Portfolio Theory and Behavioral Finance. *The Journal of Wealth Management* 7(2), 16. doi: 10.3905/jwm.2004.434562.
D'Arcangelis, A.M., and Rotundo, G. (2021). Herding in mutual funds: A complex network approach. *Journal of Business Research* 129, 679-686. doi: https://doi.org/10.1016/j.jbusres.2019.11.016.
Da, Z., Engelberg, J., and Gao, P. (2015). The Sum of All FEARS Investor Sentiment and Asset Prices. *The Review of Financial Studies* 28(1), 1-32. doi: 10.1093/rfs/hhu072.
Das, S.R., and Chen, M.Y. (2007). Yahoo! for Amazon: Sentiment Extraction from Small Talk on the Web. *Management Science* 53(9), 1375-1388. doi: 10.1287/mnsc.1070.0704.
De Bondt, W.F.M., and Thaler, R.H. (1995). "Chapter 13 Financial decision-making in markets and firms: A behavioral perspective," in *Handbooks in Operations Research and Management Science*. Elsevier), 385-410.
De Long, J.B., Shleifer, A., Summers, L.H., and Waldmann, R.J. (1990). Noise Trader Risk in Financial Markets. *Journal of Political Economy* 98(4), 703-738. doi: 10.1086/261703.
DeVault, L., Sias, R., and Starks, L. (2019). Sentiment Metrics and Investor Demand. *The Journal of Finance* 74(2), 985-1024. doi: https://doi.org/10.1111/jofi.12754.
Ding, Z., Liu, Z., Zhang, Y., and Long, R. (2017). The contagion effect of international crude oil price fluctuations on Chinese stock market investor sentiment. *Applied Energy* 187, 27-36. doi: https://doi.org/10.1016/j.apenergy.2016.11.037.

Dumitrescu, E.-I., and Hurlin, C. (2012). Testing for Granger non-causality in heterogeneous panels. *Economic Modelling* 29(4), 1450-1460. doi: https://doi.org/10.1016/j.econmod.2012.02.014.

Easley, D., and O'Hara, M. (1987). Price, trade size, and information in securities markets. *Journal of Financial Economics* 19(1), 69-90. doi: https://doi.org/10.1016/0304-405X(87)90029-8.

Edwards, W. (1982). "Conservatism in human information processing," in *Judgment under Uncertainty: Heuristics and Biases,* eds. A. Tversky, D. Kahneman & P. Slovic. (Cambridge: Cambridge University Press), 359-369.

Elton, E.J., and Gruber, M.J. (1997). Modern portfolio theory, 1950 to date. *Journal of Banking & Finance* 21(11), 1743-1759. doi: https://doi.org/10.1016/S0378-4266(97)00048-4.

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

Fama, E.F. (1998). Market efficiency, long-term returns, and behavioral finance1The comments of Brad Barber, David Hirshleifer, S.P. Kothari, Owen Lamont, Mark Mitchell, Hersh Shefrin, Robert Shiller, Rex Sinquefield, Richard Thaler, Theo Vermaelen, Robert Vishny, Ivo Welch, and a referee have been helpful. Kenneth French and Jay Ritter get special thanks.1. *Journal of Financial Economics* 49(3), 283-306. doi: https://doi.org/10.1016/S0304-405X(98)00026-9.

Fama, E.F., and French, K.R. (2004). The Capital Asset Pricing Model: Theory and Evidence. *Journal of Economic Perspectives* 18(3), 25-46. doi: 10.1257/0895330042162430.

Feldman, T. (2010). A More Predictive Index of Market Sentiment. *Journal of Behavioral Finance* 11(4), 211-223. doi: 10.1080/15427560.2010.526892.

Fisher, K.L., and Statman, M. (2000). Investor Sentiment and Stock Returns. *Financial Analysts Journal* 56(2), 16-23. doi: 10.2469/faj.v56.n2.2340.

Fisher, K.L., and Statman, M. (2003). Consumer Confidence and Stock Returns. *The Journal of Portfolio Management* 30(1), 115. doi: 10.3905/jpm.2003.319925.

Fong, W.M. (2013). Risk Preferences, Investor Sentiment and Lottery Stocks: A Stochastic Dominance Approach. *Journal of Behavioral Finance* 14(1), 42-52. doi: 10.1080/15427560.2013.759579.

Fong, W.M., and Toh, B. (2014). Investor sentiment and the MAX effect. *Journal of Banking & Finance* 46, 190-201. doi: https://doi.org/10.1016/j.jbankfin.2014.05.006.

Frugier, A. (2016). Returns, volatility and investor sentiment: Evidence from European stock markets. *Research in International Business and Finance* 38, 45-55. doi: https://doi.org/10.1016/j.ribaf.2016.03.007.

Funaoka, K., and Nishimura, Y. (2019). Private Information, Investor Sentiment, and IPO Pricing: Which Institutional Investors Are Better Informed? *Emerging Markets Finance and Trade* 55(8), 1722-1736. doi: 10.1080/1540496X.2018.1484355.

Gao, Z., Ren, H., and Zhang, B. (2020). Googling Investor Sentiment around the World. *Journal of Financial and Quantitative Analysis* 55(2), 549-580. doi: 10.1017/S0022109019000061.

Giannikos, C.I., Kakolyris, A., and Suen, T.S. (2023). Prospect theory and a manager's decision to trade a blind principal bid basket. *Global Finance Journal* 55, 100806. doi: https://doi.org/10.1016/j.gfj.2023.100806.

Gilbert, E., and Karahalios, K. (2010). Widespread Worry and the Stock Market. *Proceedings of the International AAAI Conference on Web and Social Media* 4(1), 58-65. doi: 10.1609/icwsm.v4i1.14023.

Glosten, L.R. (1989). Insider Trading, Liquidity, and the Role of the Monopolist Specialist. *The Journal of Business* 62(2), 211-235.

Golman, R., Hagmann, D., and Loewenstein, G. (2017). Information Avoidance. *Journal of Economic Literature* 55(1), 96-135. doi: 10.1257/jel.20151245.

Gu, A., and Yoo, H.I. (2021). Prospect Theory and Mutual Fund Flows. *Economics Letters* 201, 109776. doi: https://doi.org/10.1016/j.econlet.2021.109776.

Gu, Q., Li, S., Tian, S., and Wang, Y. (2023). Climate, geopolitical, and energy market risk interconnectedness: Evidence from a new climate risk index. *Finance Research Letters* 58, 104392. doi: https://doi.org/10.1016/j.frl.2023.104392.

Guo, H. (2004). A rational pricing explanation for the failure of CAPM. *Review* 86(May), 23-34.

He, Z. (2023). Geopolitical risks and investor sentiment: Causality and TVP-VAR analysis. *The North American Journal of Economics and Finance* 67, 101947. doi: https://doi.org/10.1016/j.najef.2023.101947.

Huang, D., Jiang, F., Tu, J., and Zhou, G. (2015). Investor Sentiment Aligned: A Powerful Predictor of Stock Returns. *The Review of Financial Studies* 28(3), 791-837. doi: https://doi.org/10.1093/rfs/hhu080.

Hudson, Y., Yan, M., and Zhang, D. (2020). Herd behaviour & investor sentiment: Evidence from UK mutual funds. *International Review of Financial Analysis* 71, 101494. doi: https://doi.org/10.1016/j.irfa.2020.101494.

Im, K.S., Pesaran, M.H., and Shin, Y. (2003). Testing for unit roots in heterogeneous panels. *Journal of Econometrics* 115(1), 53-74. doi: https://doi.org/10.1016/S0304-4076(03)00092-7.

Indro, D.C. (2010). "Does Mutual Fund Flow Reflect Investor Sentiment?
Handbook of Behavioral Finance." (Cheltenham, UK: Edward Elgar Publishing).

Jegadeesh, N., and Titman, S. (1993). Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency. *The Journal of Finance* 48(1), 65-91. doi: https://doi.org/10.1111/j.1540-6261.1993.tb04702.x.

Jiang, L., Wu, K., and Zhou, G. (2018). Asymmetry in Stock Comovements: An Entropy Approach. *Journal of Financial and Quantitative Analysis* 53(4), 1479-1507. doi: 10.1017/S0022109018000340.

Jolliffe, I.T., and Cadima, J. (2016). Principal component analysis: a review and recent developments. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 374(2065), 20150202. doi:

10.1098/rsta.2015.0202.
Kahneman, D., and Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica* 47(2), 263-291. doi: 10.2307/1914185.
Kao, C. (1999). Spurious regression and residual-based tests for cointegration in panel data. *Journal of Econometrics* 90(1), 1-44. doi: https://doi.org/10.1016/S0304-4076(98)00023-2.
Kassouri, Y., and Altıntaş, H. (2020). Commodity terms of trade shocks and real effective exchange rate dynamics in Africa's commodity-exporting countries. *Resources Policy* 68, 101801. doi: https://doi.org/10.1016/j.resourpol.2020.101801.
Katona, G. (1953). Rational behavior and economic behavior. *Psychological Review* 60, 307-318. doi: 10.1037/h0060640.
Keynes, J.M. (1937). The General Theory of Employment. *The Quarterly Journal of Economics* 51(2), 209-223. doi: 10.2307/1882087.
Khadjeh Nassirtoussi, A., Aghabozorgi, S., Ying Wah, T., and Ngo, D.C.L. (2014). Text mining for market prediction: A systematic review. *Expert Systems with Applications* 41(16), 7653-7670. doi: https://doi.org/10.1016/j.eswa.2014.06.009.
Kleidon, A.W. (1986). Bias in Small Sample Tests of Stock Price Rationality. *The Journal of Business* 59(2), 237-261.
Kyle, A.S. (1985). Continuous Auctions and Insider Trading. *Econometrica* 53(6), 1315-1335. doi: 10.2307/1913210.
Lee, C.M.C., Mucklow, B., and Ready, M.J. (1993). Spreads, Depths, and the Impact of Earnings Information: An Intraday Analysis. *The Review of Financial Studies* 6(2), 345-374. doi: 10.1093/rfs/6.2.345.
Lee, C.M.C., Shleifer, A., and Thaler, R.H. (1991). Investor Sentiment and the Closed-End Fund Puzzle. *The Journal of Finance* 46(1), 75-109. doi: https://doi.org/10.1111/j.1540-6261.1991.tb03746.x.
Lemmon, M., and Portniaguina, E. (2006). Consumer Confidence and Asset Prices: Some Empirical Evidence. *The Review of Financial Studies* 19(4), 1499-1529. doi: 10.1093/rfs/hhj038.
Liao, T.-L., Huang, C.-J., and Wu, C.-Y. (2011). Do fund managers herd to counter investor sentiment? *Journal of Business Research* 64(2), 207-212. doi: https://doi.org/10.1016/j.jbusres.2010.01.007.
Lintner, J. (1969). The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets: A Reply. *The Review of Economics and Statistics* 51(2), 222-224. doi: 10.2307/1926735.
Liu, S. (2015). Investor Sentiment and Stock Market Liquidity. *Journal of Behavioral Finance* 16(1), 51-67. doi: 10.1080/15427560.2015.1000334.
Ljungqvist, A., Nanda, V., and Singh, R. (2006). Hot Markets, Investor Sentiment, and IPO Pricing. *The Journal of Business* 79(4), 1667-1702. doi: 10.1086/503644.
Loughran, T.I.M., and McDonald, B. (2016). Textual Analysis in Accounting and Finance: A Survey. *Journal of Accounting Research* 54(4), 1187-1230. doi:

https://doi.org/10.1111/1475-679X.12123.
Ludvigson, S.C. (2004). Consumer Confidence and Consumer Spending. *Journal of Economic Perspectives* 18(2), 29-50. doi: 10.1257/0895330041371222.
Maddala, G.S., and Wu, S. (1999). A Comparative Study of Unit Root Tests with Panel Data and a New Simple Test. *Oxford Bulletin of Economics and Statistics* 61(S1), 631-652. doi: https://doi.org/10.1111/1468-0084.0610s1631.
Malkiel, B.G. (1977). THE VALUATION OF CLOSED-END INVESTMENT-COMPANY SHARES. *The Journal of Finance* 32(3), 847-859. doi: https://doi.org/10.1111/j.1540-6261.1977.tb01993.x.
Mao, H., Counts, S., and Bollen, J. (2011). "Predicting Financial Markets: Comparing Survey, News, Twitter and Search Engine Data". arXiv.org).
Markowitz, H. (1952). Portfolio Selection. *The Journal of Finance* 7(1), 77-91. doi: 10.2307/2975974.
Markowitz, H.M., and Todd, G.P. (2000). *Mean-variance analysis in portfolio choice and capital markets.* John Wiley & Sons.
Massa, M., and Yadav, V. (2015). Investor Sentiment and Mutual Fund Strategies. *Journal of Financial and Quantitative Analysis* 50(4), 699-727. doi: 10.1017/S0022109015000253.
Mian, G.M., and Sankaraguruswamy, S. (2012). Investor Sentiment and Stock Market Response to Earnings News. *The Accounting Review* 87(4), 1357-1384. doi: 10.2308/accr-50158.
Morris, S., Shim, I., and Shin, H.S. (2017). Redemption risk and cash hoarding by asset managers. *Journal of Monetary Economics* 89, 71-87. doi: https://doi.org/10.1016/j.jmoneco.2017.03.008.
Mossin, J. (1966). Equilibrium in a Capital Asset Market. *Econometrica* 34(4), 768-783. doi: 10.2307/1910098.
Neupane, S., Neupane, B., Paudyal, K., and Thapa, C. (2016). Domestic and foreign institutional investors' investment in IPOs. *Pacific-Basin Finance Journal* 39, 197-210. doi: https://doi.org/10.1016/j.pacfin.2016.06.011.
Nguyen, A.-N., Shahid, M.S., and Kernohan, D. (2018). Investor confidence and mutual fund performance in emerging markets. *Journal of Economic Studies* 45(6), 1288-1310. doi: 10.1108/JES-07-2017-0175.
Ofek, E., and Richardson, M. (2003). DotCom Mania: The Rise and Fall of Internet Stock Prices. *The Journal of Finance* 58(3), 1113-1137. doi: https://doi.org/10.1111/1540-6261.00560.
Pareto, V. (1909). *Manuel d'économie politique.* Giard & Brière.
Pesaran, M.H. (2007). A simple panel unit root test in the presence of cross-section dependence. *Journal of Applied Econometrics* 22(2), 265-312. doi: https://doi.org/10.1002/jae.951.
Pesaran, M.H. (2015). Testing Weak Cross-Sectional Dependence in Large Panels. *Econometric Reviews* 34(6-10), 1089-1117. doi: 10.1080/07474938.2014.956623.
Pesaran, M.H., Shin, Y., and Smith, R.J. (2001). Bounds testing approaches to the

analysis of level relationships. *Journal of Applied Econometrics* 16(3), 289-326. doi: https://doi.org/10.1002/jae.616.

Pesaran, M.H., Shin, Y., and Smith, R.P. (1999). Pooled Mean Group Estimation of Dynamic Heterogeneous Panels. *Journal of the American Statistical Association* 94(446), 621-634. doi: 10.1080/01621459.1999.10474156.

Purnanandam, A.K., and Swaminathan, B. (2004). Are IPOs Really Underpriced? *The Review of Financial Studies* 17(3), 811-848. doi: 10.1093/rfs/hhg055.

Qiu, L., and Welch, I. (2004). Investor Sentiment Measures. *National Bureau of Economic Research Working Paper Series* No. 10794. doi: 10.3386/w10794.

Ricciardi, V. (2008). "Risk: Traditional Finance versus Behavioral Finance," in *Handbook of Finance*.).

Rice, B. (2017). "The upside of the downside of modern portfolio theory", in: *Investment & Wealth Monitor, Investment Management Consultants Association.*).

Ringnér, M. (2008). What is principal component analysis? *Nature Biotechnology* 26(3), 303-304. doi: 10.1038/nbt0308-303.

Ritter, J.R. (1984). The "Hot Issue" Market of 1980. *The Journal of Business* 57(2), 215-240.

Ritter, J.R. (1991). The Long-Run Performance of initial Public Offerings. *The Journal of Finance* 46(1), 3-27. doi: https://doi.org/10.1111/j.1540-6261.1991.tb03743.x.

Ritter, J.R. (2003). Behavioral finance. *Pacific-Basin Finance Journal* 11(4), 429-437. doi: https://doi.org/10.1016/S0927-538X(03)00048-9.

Salisu, A.A., and Isah, K.O. (2017). Revisiting the oil price and stock market nexus: A nonlinear Panel ARDL approach. *Economic Modelling* 66, 258-271. doi: https://doi.org/10.1016/j.econmod.2017.07.010.

Samargandi, N. (2019). Energy intensity and its determinants in OPEC countries. *Energy* 186, 115803. doi: https://doi.org/10.1016/j.energy.2019.07.133.

Santi, C., and Zwinkels, R.C.J. (2023). Exploring style herding by mutual funds. *Journal of International Financial Markets, Institutions and Money* 85, 101762. doi: https://doi.org/10.1016/j.intfin.2023.101762.

Scheinkman, José A., and Xiong, W. (2003). Overconfidence and Speculative Bubbles. *Journal of Political Economy* 111(6), 1183-1220. doi: 10.1086/378531.

Schmeling, M. (2007). Institutional and individual sentiment: Smart money and noise trader risk? *International Journal of Forecasting* 23(1), 127-145. doi: https://doi.org/10.1016/j.ijforecast.2006.09.002.

Schmeling, M. (2009). Investor sentiment and stock returns: Some international evidence. *Journal of Empirical Finance* 16(3), 394-408. doi: https://doi.org/10.1016/j.jempfin.2009.01.002.

Sebri, M., Issoufou Ahmed, O., and Dachraoui, H. (2023). Public spending and the resource curse in WAEMU countries: An asymmetry analysis using the hidden cointegration and non-linear panel ARDL framework. *Resources Policy* 82, 103591. doi: https://doi.org/10.1016/j.resourpol.2023.103591.

Shah, T., and Baser, N. (2022). Global mutual fund market: the turn of the month effect and investment strategy. *Journal of Asset Management* 23(6), 466-476. doi: 10.1057/s41260-022-00282-0.

Sharpe, W.F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance* 19(3), 425-442. doi: https://doi.org/10.1111/j.1540-6261.1964.tb02865.x.

Shefrin, H., and Statman, M. (1985). The Disposition to Sell Winners Too Early and Ride Losers Too Long: Theory and Evidence. *The Journal of Finance* 40(3), 777-790. doi: https://doi.org/10.1111/j.1540-6261.1985.tb05002.x.

Shefrin, H., and Statman, M. (1994). Behavioral Capital Asset Pricing Theory. *Journal of Financial and Quantitative Analysis* 29(3), 323-349. doi: 10.2307/2331334.

Shiller, R.J. (1990). Speculative Prices and Popular Models. *Journal of Economic Perspectives* 4(2), 55-65. doi: 10.1257/jep.4.2.55.

Shiller, R.J. (2003). From Efficient Markets Theory to Behavioral Finance. *Journal of Economic Perspectives* 17(1), 83-104. doi: 10.1257/089533003321164967.

Shiller, R.J. (2016). *Irrational Exuberance.* Princeton University Press.

Shin, Y., and Pesaran, M.H. (1999). "An Autoregressive Distributed Lag Modelling Approach to Cointegration Analysis," in *Econometrics and Economic Theory in the 20th century,* ed. S. Strom. (Cambridge: Cambridge University Press), 371-413.

Shin, Y., Yu, B., and Greenwood-Nimmo, M. (2014). "Modelling Asymmetric Cointegration and Dynamic Multipliers in a Nonlinear ARDL Framework," in *Festschrift in Honor of Peter Schmidt: Econometric Methods and Applications,* eds. R.C. Sickles & W.C. Horrace. (New York, NY: Springer New York), 281-314.

Sias, Richard W., Starks, Laura T., and Titman, S. (2006). Changes in Institutional Ownership and Stock Returns: Assessment and Methodology. *The Journal of Business* 79(6), 2869-2910. doi: 10.1086/508002.

Sibley, S.E., Wang, Y., Xing, Y., and Zhang, X. (2016). The information content of the sentiment index. *Journal of Banking & Finance* 62, 164-179. doi: https://doi.org/10.1016/j.jbankfin.2015.10.001.

Simon, H.A. (1955). A Behavioral Model of Rational Choice. *The Quarterly Journal of Economics* 69(1), 99-118. doi: 10.2307/1884852.

Singhvi, V. (2001). *Investor sentiment: Its measurement and dimensions.* New York University, Graduate School of Business Administration.

Smales, L.A. (2017). The importance of fear: investor sentiment and stock market returns. *Applied Economics* 49(34), 3395-3421. doi: 10.1080/00036846.2016.1259754.

Smith, A. (2010a). *The theory of moral sentiments.* Penguin.

Smith, A. (2010b). *The Wealth of Nations: An inquiry into the nature and causes of the Wealth of Nations.* Harriman House Limited.

Stambaugh, R.F., Yu, J., and Yuan, Y. (2012). The short of it: Investor sentiment and anomalies. *Journal of Financial Economics* 104(2), 288-302. doi: https://doi.org/10.1016/j.jfineco.2011.12.001.

Stambaugh, R.F., Yu, J., and Yuan, Y.U. (2015). Arbitrage Asymmetry and the Idiosyncratic Volatility Puzzle. *The Journal of Finance* 70(5), 1903-1948. doi: https://doi.org/10.1111/jofi.12286.

Suanin, W. (2021). Demand Elasticity of Processed Food Exports from Developing Countries: A Panel Analysis of US Imports. *Journal of Agricultural Economics* 72(2), 413-429. doi: https://doi.org/10.1111/1477-9552.12409.

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. doi: https://doi.org/10.1111/j.1540-6261.2007.01232.x.

Thaler, R.H. (2016). Behavioral Economics: Past, Present, and Future. *American Economic Review* 106(7), 1577-1600. doi: 10.1257/aer.106.7.1577.

Tobin, J. (1958). Liquidity Preference as Behavior Towards Risk. *The Review of Economic Studies* 25(2), 65-86. doi: 10.2307/2296205.

Tomer, J.F. (2007). What is behavioral economics? *The Journal of Socio-Economics* 36(3), 463-479. doi: https://doi.org/10.1016/j.socec.2006.12.007.

Treynor, J.L. (1961). Market value, time, and risk. *Time, and Risk (August 8, 1961)*.

Tversky, A., and Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases. *Science* 185(4157), 1124-1131. doi: 10.1126/science.185.4157.1124.

Ugurlu-Yildirim, E., Kocaarslan, B., and Ordu-Akkaya, B.M. (2021). Monetary policy uncertainty, investor sentiment, and US stock market performance: New evidence from nonlinear cointegration analysis. *International Journal of Finance & Economics* 26(2), 1724-1738. doi: https://doi.org/10.1002/ijfe.1874.

Wang, C., and Han, J. (2023). Prospect theory and mutual fund flows: Evidence from China. *Pacific-Basin Finance Journal* 80, 102067. doi: https://doi.org/10.1016/j.pacfin.2023.102067.

Wang, W.-Z., Liu, L.-C., Liao, H., and Wei, Y.-M. (2021). Impacts of urbanization on carbon emissions: An empirical analysis from OECD countries. *Energy Policy* 151, 112171. doi: https://doi.org/10.1016/j.enpol.2021.112171.

Wang, W. (2020). Institutional investor sentiment, beta, and stock returns. *Finance Research Letters* 37, 101374. doi: https://doi.org/10.1016/j.frl.2019.101374.

Warther, V.A. (1995). Aggregate mutual fund flows and security returns. *Journal of Financial Economics* 39(2), 209-235. doi: https://doi.org/10.1016/0304-405X(95)00827-2.

Weiss, K. (1989). The Post-Offering Price Performance of Closed-End Funds. *Financial Management* 18(3), 57-67. doi: 10.2307/3665649.

Wermers, R. (1999). Mutual Fund Herding and the Impact on Stock Prices. *The Journal of Finance* 54(2), 581-622. doi: https://doi.org/10.1111/0022-1082.00118.

Yu, B., Shen, Y., Jin, X., and Xu, Q. (2022). Does prospect theory explain mutual fund performance? Evidence from China. *Pacific-Basin Finance Journal* 73, 101766. doi: https://doi.org/10.1016/j.pacfin.2022.101766.

Yu, J., and Yuan, Y. (2011). Investor sentiment and the mean–variance relation. *Journal of Financial Economics* 100(2), 367-381. doi: https://doi.org/10.1016/j.jfineco.2010.10.011.

Zaremba, A., Cakici, N., Demir, E., and Long, H. (2022). When bad news is good news: Geopolitical risk and the cross-section of emerging market stock returns. *Journal of Financial Stability* 58, 100964. doi: https://doi.org/10.1016/j.jfs.2021.100964.

Zhang, S., Farzan, Y., Huy, P., and Muhammad, W. (2021). Do Transparency and Anti-Monopoly Policies Matter for Financial Development? Evidence from a Panel ARDL-PMG Approach. *Journal of Applied Economics* 24(1), 1-16. doi: 10.1080/15140326.2020.1838113.

Zhang, X., Fuehres, H., and Gloor, P.A. (2011). Predicting Stock Market Indicators Through Twitter "I hope it is not as bad as I fear". *Procedia - Social and Behavioral Sciences* 26, 55-62. doi: https://doi.org/10.1016/j.sbspro.2011.10.562.

Zhengke, Y., Chunyan, H., Linjie, H., Guangda, O., and Fenghua, W. (2020). The Dynamic Time-frequency Relationship between International Oil Prices and Investor Sentiment in China: A Wavelet Coherence Analysis. *Energy Journal* 41(5), 251-270. doi: 10.5547/01956574.41.5.fwen.

Zweig, M.E. (1973). An Investor Expectations Stock Price Predictive Model Using Closed-End Fund Premiums. *The Journal of Finance* 28(1), 67-78. doi: 10.2307/2978169.

Zwick, W.R., and Velicer, W.F. (1986). Comparison of five rules for determining the number of components to retain. *Psychological Bulletin* 99(3), 432-442. doi: 10.1037/0033-2909.99.3.432.

# Appendix

### Table A.X.2. Bounds test for long-run cointegration

| | | CCI | GO | BW | NIPO | RIPO | VIX | SENT |
|---|---|---|---|---|---|---|---|---|
| Ps | F-statistic | 24.525 | 5.370 | 3.836 | 13.988 | 18.671 | 11.343 | 11.833 |
| Pc | F-statistic | 8.438 | 9.059 | 5.566 | 4.530 | 5.760 | 6.147 | 13.422 |
| Upper bound and lower bound | | | 10% | | 5% | | 1% | |
| | | | I (0) | I (1) | I (0) | I (1) | I (0) | I (1) |
| Critical value | | | 2.450 | 3.520 | 2.860 | 4.010 | 3.740 | 5.060 |

**Notes:** Ps and Pc refer to the percentage of capital amount of mutual funds invested in common stock, and cash, respectively. CCI means the consumer confidence index, GO means google trend searching volume of financial crisis and economic recession, BW indicates Baker and Wurgler index, NIPO means the number of IPOs, RIPO means the first day return of IPOs, VIX means the volatility index. The table demonstrates the bounds test for panel ARDL. The common F-statistics are obtained by the Wald test and then compared with the critical value. If the value is above the upper bound I (1), which suggests that there is long-run cointegration, if the F-statistic is below the lower bound I (0), then there is no long-run cointegration. However, if the value falls between the upper bound and the lower bound, then it indicates that the long-cointegration is inconclusive.

### Table A.X.2. Correlation matrix

| | ***Ps*** | ***Pc*** | ***CCI*** | ***GO*** | ***BW*** | ***NIPO*** | ***RIPO*** | ***VIX*** | ***IR*** | ***UEMP*** | ***GPR*** |
|---|---|---|---|---|---|---|---|---|---|---|---|
| ***Ps*** | 1.000 | | | | | | | | | | |
| ***Pc*** | -0.101 | 1.000 | | | | | | | | | |
| ***CCI*** | -0.033 | -0.169 | 1.000 | | | | | | | | |
| ***GO*** | 0.006 | 0.154 | -0.567 | 1.000 | | | | | | | |
| ***BW*** | 0.018 | -0.022 | 0.396 | -0.446 | 1.000 | | | | | | |
| ***NIPO*** | -0.015 | -0.053 | 0.235 | -0.329 | 0.688 | 1.000 | | | | | |
| ***RIPO*** | -0.017 | -0.082 | 0.290 | -0.444 | 0.486 | 0.572 | 1.000 | | | | |
| ***VIX*** | 0.062 | 0.156 | -0.511 | 0.692 | -0.112 | -0.068 | -0.047 | 1.000 | | | |
| ***IR*** | 0.124 | 0.071 | 0.286 | -0.006 | 0.149 | -0.203 | -0.153 | 0.130 | 1.000 | | |
| ***UEMP*** | -0.004 | 0.058 | -0.781 | 0.331 | -0.326 | -0.022 | 0.052 | 0.397 | -0.520 | 1.000 | |
| ***GPR*** | -0.013 | -0.038 | 0.233 | -0.167 | -0.154 | -0.320 | -0.268 | -0.328 | -0.006 | -0.279 | 1.000 |

### Table A.X.2. Mutual funds identifier from CRSP database

| | | | | |
|---|---|---|---|---|
| 105 | 43541 | 45842 | 45974 | 46137 |
| 2746 | 43573 | 45862 | 45977 | 46139 |

| | | | | |
|---|---|---|---|---|
| 29469 | 43624 | 45870 | 45979 | 46140 |
| 29483 | 45723 | 45875 | 45982 | 46142 |
| 29493 | 45734 | 45879 | 45990 | 46194 |
| 29531 | 45735 | 45884 | 46016 | 46206 |
| 29575 | 45736 | 45889 | 46018 | 46225 |
| 29576 | 45737 | 45895 | 46022 | 46256 |
| 29582 | 45738 | 45900 | 46025 | 46257 |
| 29586 | 45740 | 45905 | 46042 | 46287 |
| 30649 | 45741 | 45910 | 46043 | 46313 |
| 31750 | 45750 | 45915 | 46044 | 46314 |
| 31864 | 45774 | 45920 | 46047 | 46323 |
| 43521 | 45778 | 45924 | 46058 | 46330 |
| 43532 | 45808 | 45925 | 46061 | 46335 |
| 43533 | 45809 | 45926 | 46105 | 46343 |
| 43534 | 45833 | 45927 | 46115 | 46410 |
| 43537 | 45836 | 45944 | 46129 | 46411 |
| 43538 | 45839 | 45969 | 46131 | 46466 |

**Figure A.X.1 Scree plot of second-stage principal analysis**

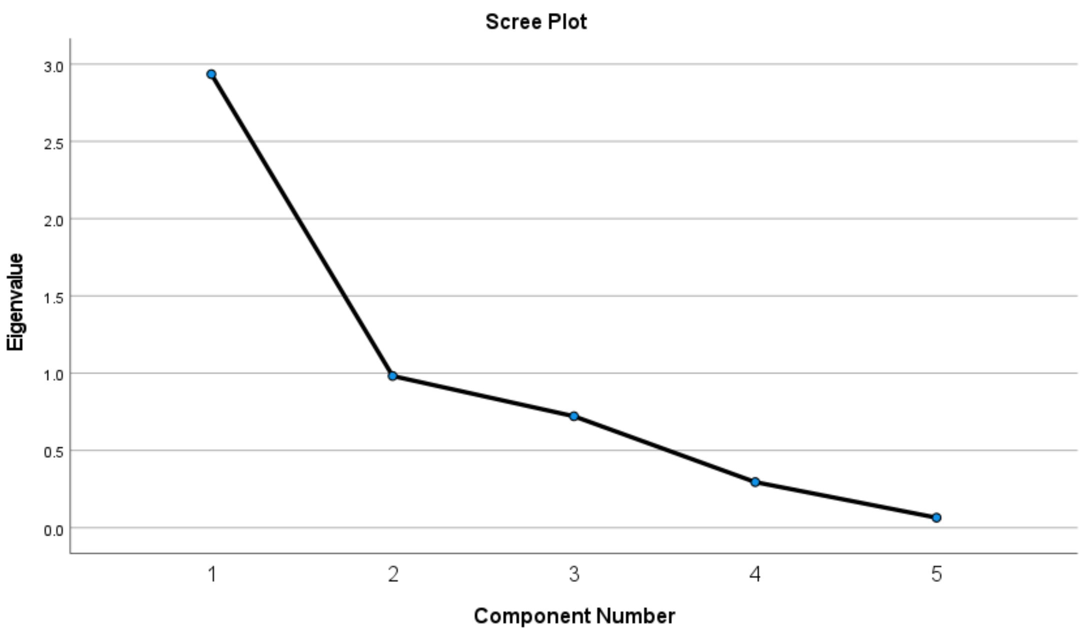
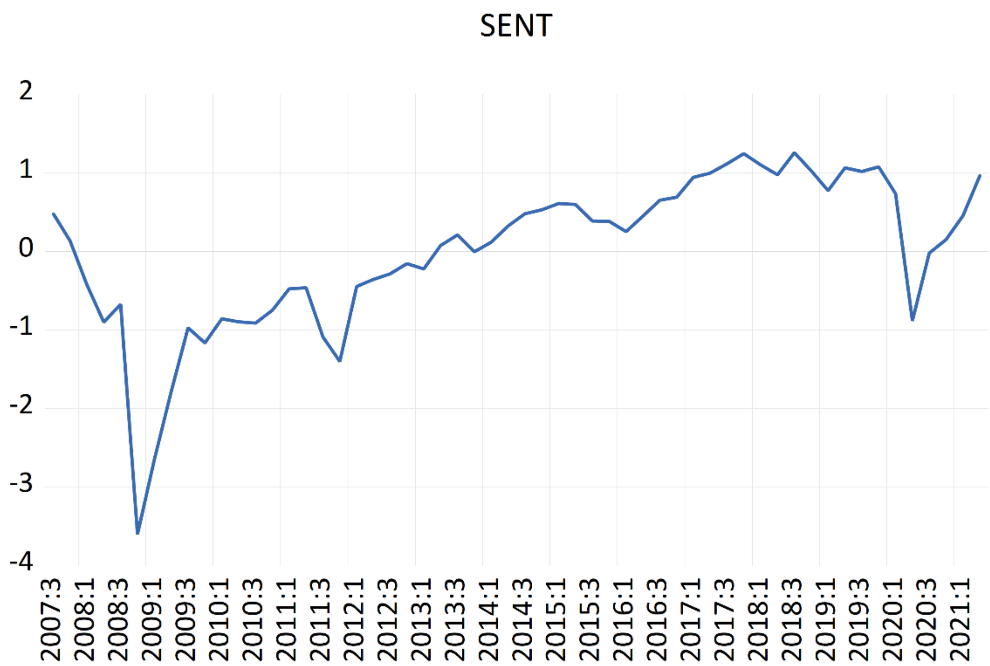



**Figure A.X.2 SENT index**




Robustness check with new control variables FRFC (Refinitiv US financials price return index percentage change of each quarter) and SP5TER (sp500 total return index percentage change of each quarter). These 2 variables will control the fluctuation of stock market price. Table A.X.3 shows CCI has a positive effect on Ps in the long run, implying that higher consumer confidence encourages greater investment in common stocks. GO appears to have a minor and non-significant relationship. BW index, a composite indicator, has a strong positive influence, suggesting that increased sentiment drives higher stock investments. NIPO and RIPO both positively influence Ps, indicating that IPO activity signals opportunities that attract fund managers towards stocks market. VIX shows a positive effect, possibly indicating a strategic positioning for potential returns during periods of uncertainty.

Table. A.X.4 shows that CCI negatively affects cash holdings, implying that increased consumer confidence prompts managers to reduce cash and allocate more to stocks.

GO shows a positive relationship with cash, which may indicate a risk-averse attitude during period with high uncertainty. BW index also increases cash holdings, suggesting some hedging behaviour despite generally positive market outlooks. As for NIPO, RIPO and VIX, none of them have significant impact on mutual funds cash holdings. Table. A.X.5 demonstrates that the SENT index positively impacts Ps and negatively impacts Pc in the long run, meaning that heightened sentiment drives more investment in stocks while reducing cash reserves.

In summary, investor sentiment still plays a significant role in determining mutual fund portfolio composition after control the price fluctuation of stock market, with higher sentiment generally leading to greater investment in common stocks and a reduction in cash reserves.

**Table A.X.3. Ps Robustness check with FRFC and SP5TER**

| | | | ***Ps*** | | | |
|---|---|---|---|---|---|---|
| ***Variables*** | **(a)** ***Survey* base** | **(b)** ***Internet base*** | **(c)** ***Composite*** | | **(d)** ***Market base*** | |
| | | | ***Long run equation*** | | | |
| ***CCI*** | 0.046* (0.007) | | | | | |
| ***GO*** | | -0.004 (0.011) | | | | |
| ***BW*** | | | 0.769* (0.324) | | | |
| ***NIPO*** | | | | 0.008* (0.002) | | |
| ***RIPO*** | | | | | 0.075* (0.011) | |
| ***VIX*** | | | | | | 0.103* (0.020) |
| ***IR*** | 0.596* (0.118) | 0.482* (0.124) | 0.495* (0.116) | 0.578* (0.119) | 0.789* (0.126) | 0.401* (0.114) |
| ***UEMP*** | 0.574* (0.108) | 0.002 (0.067) | 0.036 (0.064) | 0.025 (0.058) | 0.124* (0.063) | -0.198* (0.069) |
| ***GPR*** | 0.669* (0.282) | 0.225 (0.281) | 0.328 (0.298) | 0.624* (0.310) | 1.054* (0.319) | 0.612* (0.291) |
| ***FRFC*** | 14.297* (2.505) | 8.460* (2.534) | 7.910* (2.412) | 5.214* (2.345) | 10.044* (2.494) | 5.507* (2.346) |
| ***SP5TR*** | -14.041* (3.876) | -0.063 (3.964) | 1.482 (3.249) | 2.418 (3.204) | -4.764 (3.544) | 10.972* (3.780) |
| | | | ***Short run equation*** | | | |
| ***COINTEQ*** | -0.336* (0.023) | -0.331* (0.022) | -0.336* (0.022) | -0.333* (0.022) | -0.334* (0.022) | -0.333* (0.022) |
| ***D(CCI)*** | -0.0003 (0.008) | | | | | |
| ***D(GO)*** | | -0.006 (0.013) | | | | |
| ***D(BW)*** | | | -0.086 (0.411) | | | |
| ***D(NIPO)*** | | | | 0.0004 (0.002) | | |
| ***D(RIPO)*** | | | | | -0.019* (0.007) | |
| ***D(VIX)*** | | | | | | -0.022 |

| | | | | | | |
|---|---|---|---|---|---|---|
| | | | | | | (0.013) |
| ***D(IR)*** | -1.717* | -1.516* | -1.547* | -1.502* | -1.470* | -1.354* |
| | (0.579) | (0.597) | (0.564) | (0.562) | (0.574) | (0.564) |
| ***D(UEMP)*** | -0.256* | -0.127 | -0.133 | -0.124 | -0.136 | -0.087 |
| | (0.082) | (0.597) | (0.079) | (0.086) | (0.082) | (0.083) |
| ***D(GPR)*** | -0.138 | -0.073 | -0.071 | -0.091 | -0.208 | -0.145 |
| | (0.122) | (0.135) | (0.126) | (0.126) | (0.129) | (0.127) |
| ***D(FRFC)*** | -5.318* | -4.597* | -4.324* | -4.021* | -5.006* | -4.112* |
| | (1.111) | (1.158) | (1.086) | (1.161) | (1.084) | (1.119) |
| ***D(SP5TR)*** | 6.144* | 3.725* | 3.338 | 3.433 | 4.867* | 1.500 |
| | (1.742) | (1.618) | (1.732) | (1.816) | (1.676) | (0.157) |
| ***C*** | 25.292* | 27.715* | 28.007* | 27.312* | 26.428* | 27.215* |
| | (2.019) | (2.115) | (2. 086) | (2.105) | (2.017) | (2.071) |

Notes: * denotes statistically significant at 5% level. The results demonstrate coefficient and standard error with parentheses.

# Table A.X.4. Pc Robustness check with FRFC and SP5TER

| | | | ***Pc*** | | | |
|---|---|---|---|---|---|---|
| ***Variables*** | **(a)** *Survey* **base** | **(b)** ***Internet base*** | **(c)** ***Composite*** | | **(d)** ***Market base*** | |
| | | | ***Long run equation*** | | | |
| ***CCI*** | -0.002* | | | | | |
| | (0.001) | | | | | |
| ***GO*** | | 0.004* | | | | |
| | | (0.001) | | | | |
| ***BW*** | | | 0.063* | | | |
| | | | (0.032) | | | |
| ***NIPO*** | | | | 0.0001 | | |
| | | | | (0.0001) | | |
| ***RIPO*** | | | | | -0.001 | |
| | | | | | (0.001) | |
| ***VIX*** | | | | | | 0.003 |
| | | | | | | (0.002) |
| ***IR*** | 0.021 | 0.035* | 0.042* | 0.041* | 0.032* | 0.030* |
| | (0.012) | (0.011) | (0.011) | (0.012) | (0.012) | (0.012) |
| ***UEMP*** | -0.020 | 0.005 | 0.016* | 0.013* | 0.008 | 0.004 |
| | (0.011) | (0.006) | (0.006) | (0.006) | (0.006) | (0.006) |
| ***GPR*** | -0.010 | 0.027 | 0.028 | 0.026 | 0.004 | 0.022 |
| | (0.030) | (0.025) | (0.029) | (0.032) | (0.032) | (0.030) |
| ***FRFC*** | -0.091 | 0.214 | 0.370 | 0.271 | 0.206 | 0.214 |
| | (0.264) | (0.233) | (0.232) | (0.246) | (0.255) | (0.244) |
| ***SP5TR*** | -0.034 | -0.588 | -0.803* | -0.786* | -0.661 | -0.819* |
| | (0.411) | (0.366) | (0.317) | (0.339) | (0.366) | (0.393) |
| | | | ***Short run equation*** | | | |
| ***COINTEQ*** | -0.563* | -0.568* | -0.569* | -0.554* | -0.558* | -0.559* |
| | (0.028) | (0.028) | (0.028) | (0.028) | (0.028) | (0.027) |
| ***D(CCI)*** | -0.037* | | | | | |
| | (0.012) | | | | | |
| ***D(GO)*** | | 0.049* | | | | |
| | | (0.014) | | | | |
| ***D(BW)*** | | | -0.760* | | | |
| | | | (0.231) | | | |
| ***D(NIPO)*** | | | | -0.001* | | |
| | | | | (0.001) | | |
| ***D(RIPO)*** | | | | | -0.010* | |
| | | | | | (0.005) | |
| ***D(VIX)*** | | | | | | 0.033* |
| | | | | | | (0.011) |
| ***D(IR)*** | -0.154 | -0.866* | -0.425* | -0.392* | -0.424* | -0.463* |
| | (0.241) | (0.156) | (0.195) | (0.197) | (0.197) | (0.182) |
| ***D(UEMP)*** | -0.081* | -0.097* | -0.016 | 0.002 | 0.012 | -0.053 |
| | (0.034) | (0.032) | (0.042) | (0.044) | (0.044) | (0.037) |
| ***D(GPR)*** | 0.392* | 0.311* | 0.148 | 0.111 | 0.118 | 0.178 |
| | (0.109) | (0.107) | (0.125) | (0.098) | (0.101) | (0.100) |
| ***D(FRFC)*** | 1.978 | 1.634 | 1.249 | 1.393 | 1.298 | 1.462 |
| | (1.070) | (1.044) | (1.013) | (1.049) | (1.028) | (1.036) |
| ***D(SP5TR)*** | -3.745* | -2.067 | -2.890* | -3.057* | -2.879* | -1.483 |
| | (1.569) | (1.268) | (1.453) | (1.488) | (1.455) | (1.212) |
| ***C*** | 1.104* | 0.755* | 0.795* | 0.767* | 0.834* | 0.781* |
| | (0.108) | (0.107) | (0.107) | (0.104) | (0.104) | (0.104) |

Notes: * denotes statistically significant at 5% level. The results demonstrate coefficient and standard error with parentheses.

**Table A.X.5. Pc Robustness check with FRFC and SP5TER**

| Variables | *Ps* | *Pc* |
|---|---|---|
| | ***Long run equation*** | |
| ***SENT*** | $0.508^{*}$ | $-0.082^{*}$ |
| | (0.217) | (0.020) |
| ***IR*** | $0.536^{*}$ | 0.014 |
| | (0.130) | (0.012) |
| ***UEMP*** | 0.175 | $-0.023^{*}$ |
| | (0.110) | (0.026) |
| ***GPR*** | 0.238 | 0.014 |
| | (0.287) | (0.024) |
| ***FRFC*** | 9.999 | -0.055 |
| | (2.591) | (0.236) |
| ***SP5TR*** | -2.867 | -0.216 |
| | (4.369) | (0.400) |
| | ***Short run equation*** | |
| ***COINTEQ*** | $-0.329^{*}$ | $-0.564^{*}$ |
| | (0.022) | (0.028) |
| ***D(SENT)*** | -0.131 | $-0.481^{*}$ |
| | (0.165) | (0.151) |
| ***D(IR)*** | $-1.677^{*}$ | $-0.492^{*}$ |
| | (0.587) | (0.181) |
| ***D(UEMP)*** | -0.186 | $-0.100^{*}$ |
| | (0.104) | (0.036) |
| ***D(GPR)*** | -0.029 | $0.272^{*}$ |
| | (0.131) | (0.101) |
| ***D(FRFC)*** | $-4.601^{*}$ | $2.386^{*}$ |
| | (1.195) | (2.085) |
| ***D(SP5TR)*** | $4.205^{*}$ | $-3.686^{*}$ |
| | (1.788) | (1.556) |
| ***C*** | $27.062^{*}$ | $0.943^{*}$ |
| | (2.250) | (0.106) |

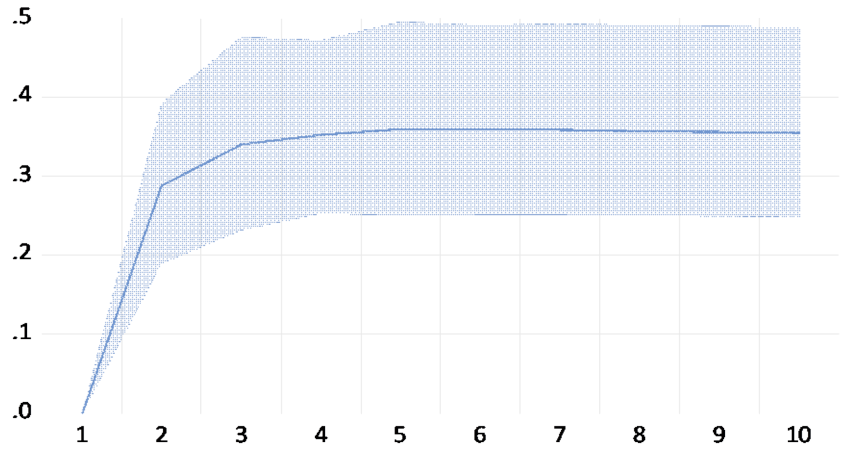
Notes: * denotes statistically significant at 5% level. The results demonstrate coefficient and standard error with parentheses. Ps and Pc refer to the percentage of capital amount of mutual funds invested in common stock, and cash, respectively. SENT means the principal component index, IR means short-term interest rate, UEMP indicates the unemployment rate, and GPR refers to the geopolitical risk index.

The graphs below are VAR results with new control variables FRFC (Refinitiv US financials price return index percentage change of each quarter) and SP5TER (sp500 total return index percentage change of each quarter). PCA1 is the SENT composite index introduced in chapter 5. Ps is stock investment percentage and Pc is cash holding percentage.

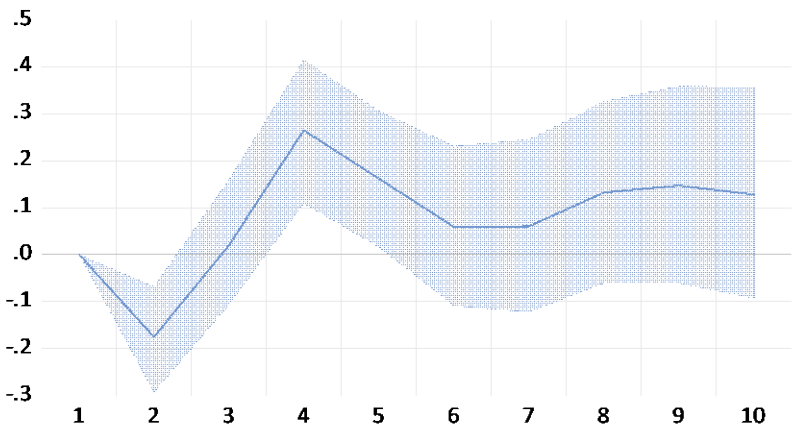
The VAR results suggest that the 95% confidence intervals indicate the periods during which the responses are statistically significant. If the shaded area includes zero, the response may not be significant. And all sentiment indices have significant positive effect on Ps. In addition, VAR results also indicate that the RIPO, BW and VIX have no significant effect on Pc, while CCI, GO and SENT have significant effect. The NIPO will have significant short run effect on Pc. The results consistent with the regression results presented in Table A.X.3, Table. A.X.4 and Table A.X.5.

Panel VAR results

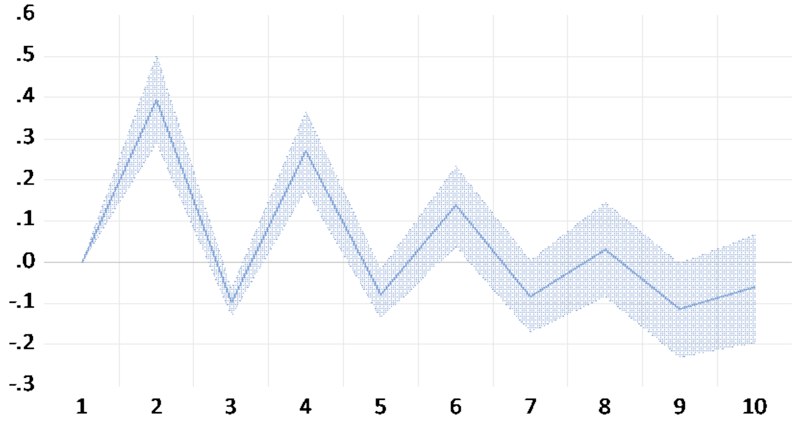
Response of PS to D(CCI) Cholesky One S.D. (d.f. adjusted) Innovation
95% CI using Standard percentile bootstrap with 500 bootstrap repetitions



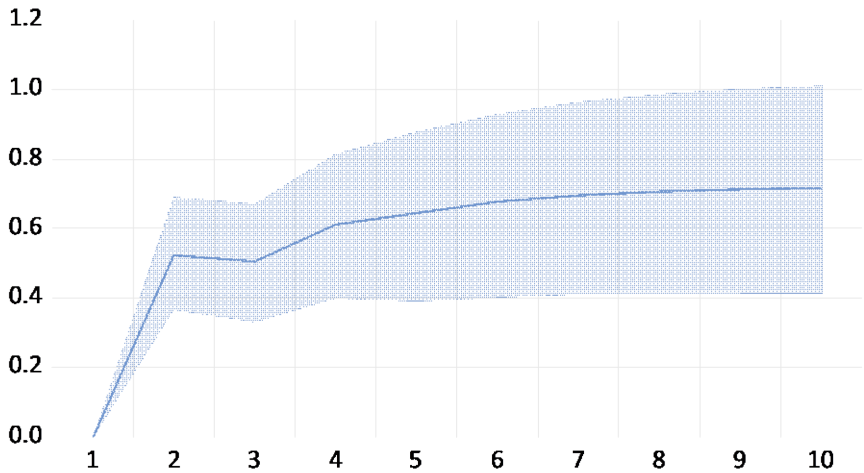

Response of PS to GO Cholesky One S.D. (d.f. adjusted) Innovation
95% CI using Standard percentile bootstrap with 500 bootstrap repetitions



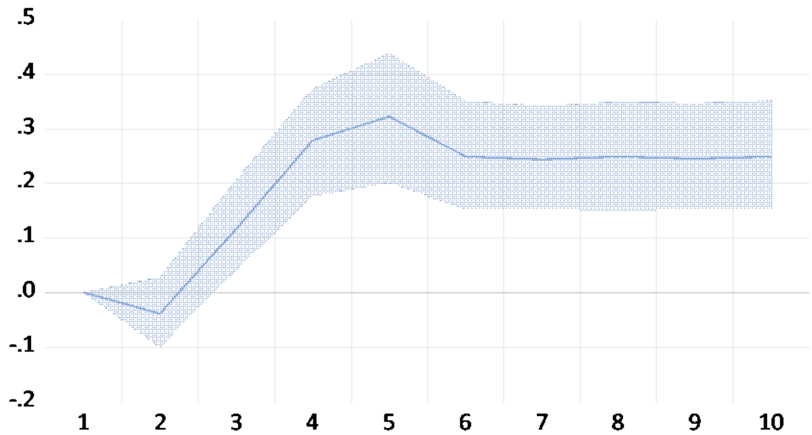
Response of PS to NIPO Cholesky One S.D. (d.f. adjusted) Innovation
95% CI using Standard percentile bootstrap with 500 bootstrap repetitions



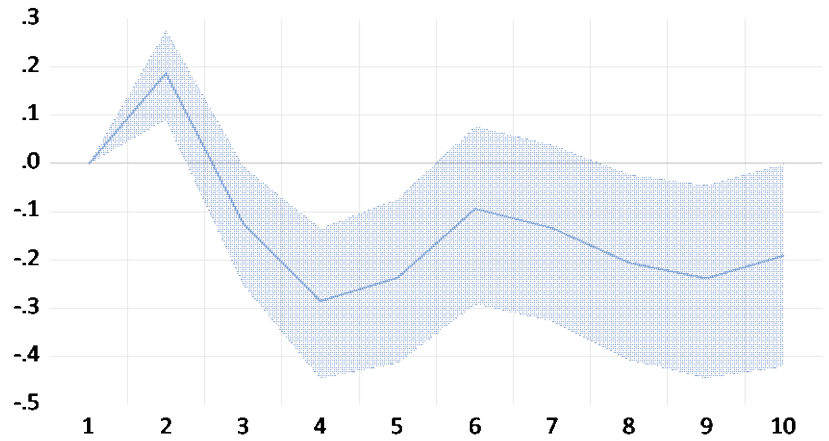

Response of PS to RIPO Cholesky One S.D. (d.f. adjusted) Innovation
95% CI using Standard percentile bootstrap with 500 bootstrap repetitions



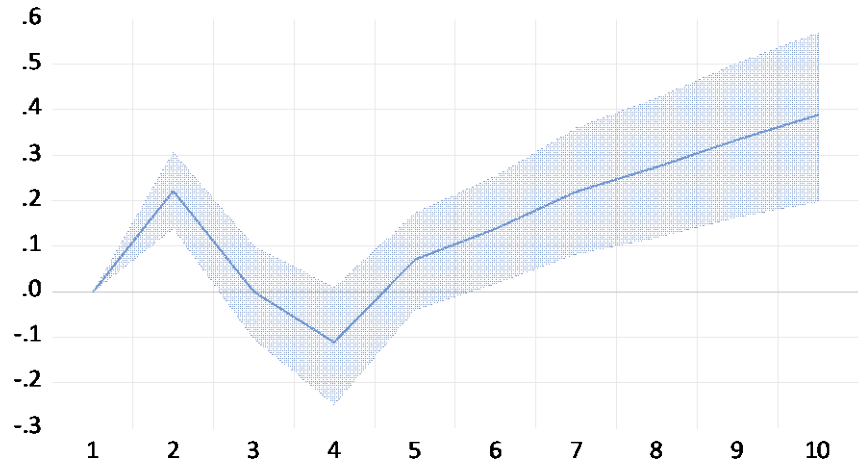

Response of PS to D(VIX) Cholesky One S.D. (d.f. adjusted) Innovation
95% CI using Standard percentile bootstrap with 500 bootstrap repetitions



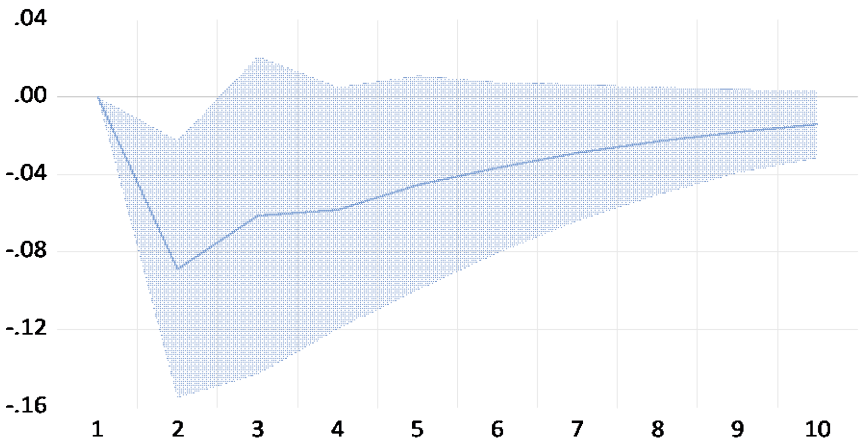
Response of PS to D(BW) Cholesky One S.D. (d.f. adjusted) Innovation
95% CI using Standard percentile bootstrap with 500 bootstrap repetitions



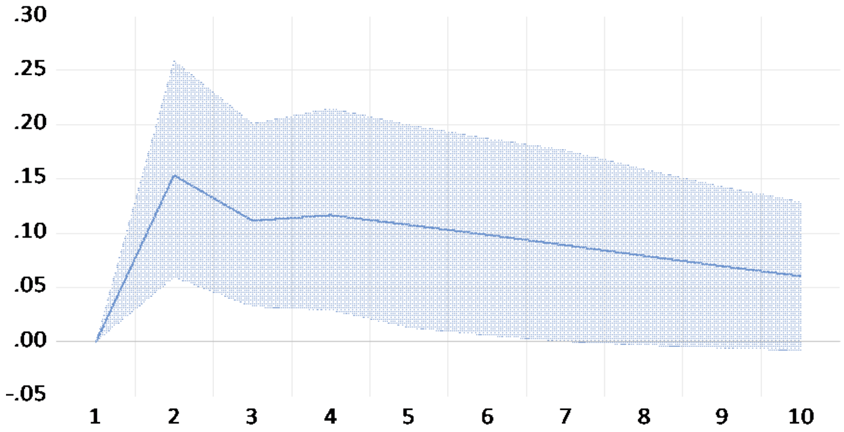
Response of PS to PCA1 Cholesky One S.D. (d.f. adjusted) Innovation
95% CI using Standard percentile bootstrap with 500 bootstrap repetitions



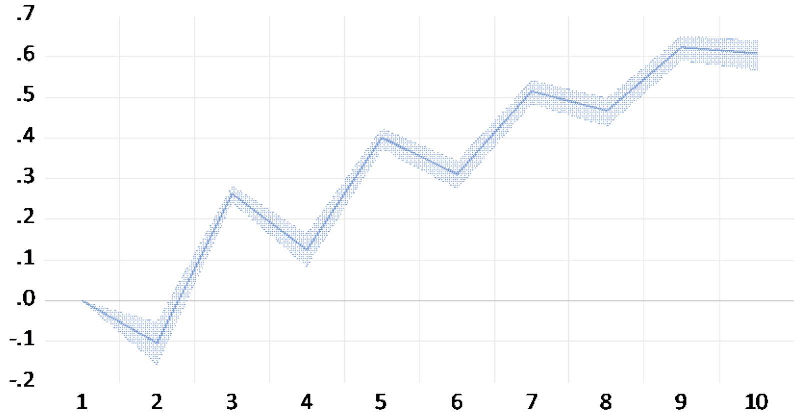
Response of PC to D(CCI) Cholesky One S.D. (d.f. adjusted) Innovation
95% CI using Standard percentile bootstrap with 500 bootstrap repetitions



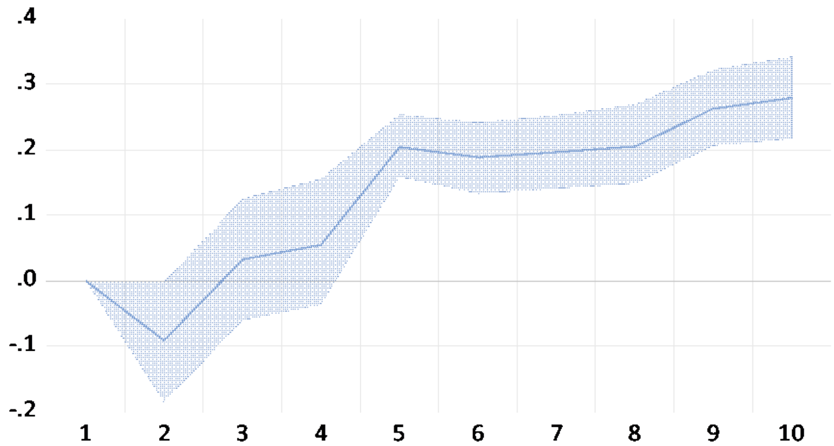
Response of PC to GO Cholesky One S.D. (d.f. adjusted) Innovation
95% CI using Standard percentile bootstrap with 500 bootstrap repetitions



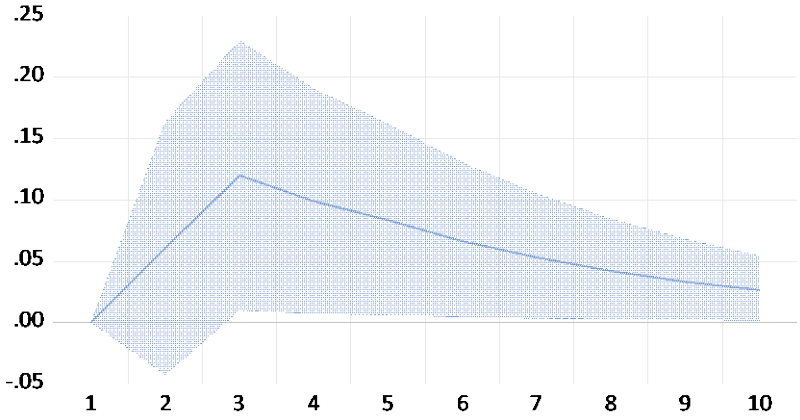

Response of PC to NIPO Cholesky One S.D. (d.f. adjusted) Innovation
95% CI using Standard percentile bootstrap with 500 bootstrap repetitions



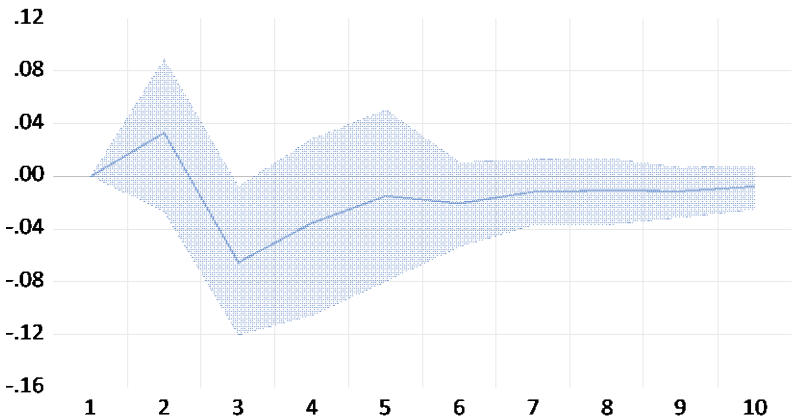

Response of PC to RIPO Cholesky One S.D. (d.f. adjusted) Innovation
95% CI using Standard percentile bootstrap with 500 bootstrap repetitions



Response of PC to D(BW) Cholesky One S.D. (d.f. adjusted) Innovation
95% CI using Standard percentile bootstrap with 500 bootstrap repetitions




Response of PC to D(VIX) Cholesky One S.D. (d.f. adjusted) Innovation
95% CI using Standard percentile bootstrap with 500 bootstrap repetitions




Response of PC to PCA1 Cholesky One S.D. (d.f. adjusted) Innovation
95% CI using Standard percentile bootstrap with 500 bootstrap repetitions